AUC Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 16 2021

Amsterdam university college
AUC
UNIVERSITY OF AMSTERDAM
imprint.
VU UNIVERSITY AMSTERDAM

Amsterdam University College Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 16 2021

Published by
inprint.

AUC Undergraduate Journal of Liberal Arts & Sciences
Capstone Issue Vol. 16 2021, published by InPrint

The Amsterdam University College (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 composed of members of the InPrint board, a registered AUCSA committee.

Foreword

It is my great pleasure and privilege to introduce the 16th volume of Amsterdam University Colleges's (AUC) Undergraduate Journal of Liberal Arts and Sciences.

The articles in this Issue are selected from AUC students' Capstone thesis - the final independent research undertaken by the undergraduate students. With three distinct departments: Sciences, Social Sciences, and Humanities, the Issue presents two of the most remarkable Capstones from each department. This Issue begins with Gesa Mueller's detailed primary research into the diel variation of macroinvertebrate assemblages in the coasts of Greece. Following that, Karla M. Rojas's compelling discussion of the recently-discovered galaxy NGC1052–DF2, and its implications in re-sparking the modified gravity debate. Then, Carolina Resigotti's illuminating thesis on the role of Rwandan youth in the diffusion of reconciliation norms. Next, Salomé Petit-Siemen's fascinating exploration into the socio-political implications of algorithmic power in social media. Following this, Lisa Philippo's ingenious paper that deconstructs the gentrification of online sexuality through analyzing OnlyFans. And finally, Miglė Gerčaitė's intricate case of counter-monumentality in Soviet Lithuania. While the research focus runs the gamut, all these papers are all tied together by their depth of critical analysis and innovation of thought.

The Issue would not be possible without the efforts of all the editors, who rigorously collaborate with the authors to continually raise the bar of excellence in the Issue. Special thanks go to the Head Editors of each department, Aada Kallio, Céline Paré, and Casey Ansara for their meticulous and tireless contributions. I would also like to express heartfelt thanks to the authors for their continued engagement and revision of their research. Finally, I would like to thank the continual support of all the AUC staff, specifically Wade Geary the Capstone Coordinator, and the Academic Writing Skills and Advanced Research Writing teams, whose support we are grateful for.

I am humbled by the level of academic excellence in this Issue. It is my sincere hope that you, the reader, will come away from this Issue inspired by the thought-provoking analysis presented by each of the theses, and perhaps feel inclined to contribute to our next Issue!

Aditi Rai Sia, Editor-in-Chief

A note from the photographers

InPrint has a long-standing collaboration with RAW to take cover photographs for each paper. The abstracts of all six papers published in this issue were sent to the photographers as inspiration for their photographs. The captions below give a short explanation of the artists' thought process and interpretation of their work.

***Diel Fade* by Stef Deuring** for Gesa Mueller's *A quantitative analysis of diel variation in macroinvertebrate assemblages in the coastal waters of Lipsi island, Greece.* In order to visually represent the night and day cycle, the picture fades from the dark blue of murky nocturnal waters to the brightness of full daylight. The macroinvertebrates loom larger than they are in real life, trying to visualise their often underestimated importance for the marine environment.

***Homemade Galactic Development* by Iacob Postavaru** for Karla M. Rojas's *NGC 1052–DF2: The galaxy re-sparking the modified gravity debate.* This diffuse galaxy lacks more than half of its expected dark matter. This means that the galaxy would not have formed through the traditional hierarchical process of galaxy formation. However, this is very different from NGC1052–DF2's case since its dynamics can be explained without studying outer space but instead with the fluid mechanics of oil and water.

***Beyond Here* by Richard Essink** for Carolina Resigotti's *Young Entrepreneurs of Reconciliation.* The burned out husk of a tree represents the collective trauma of the Rwandan Genocide. Candles burn in mourning vigil and collective reappraisal of this tragic event. Finally, leaving the desolateness of this tragedy behind, the young entrepreneur climbs upwards, towards hope.

***Erased* by Richard Essink** for Salomé Petit-Siemens's *Social Media Platforms and a New Regime of (In)visibility.* Blurred, stuck between this reality and their digital presence, the figure in the picture loses their identity. Unidentifiable, maligned, he is supposed to represent the potential of social media to effectively erase the voices of suppressed communities.

***Moulin Rouge* by Daria Roman** for Lisa Philippo's *Digital Whores Doing Pay-Per-View Chores.* Elements that are commonly associated with sex workers, seduction, and aestheticism are combined. In order to achieve a mysterious, seductive aesthetic, decisions about the lighting, exposure, and composition were made. I incorporated a deep red light, minimized the exposure, and decided to only focus on a part of the leg being touched by the hand in a sensual manner instead of photographing someone and including a face. I believe this helped achieve a picture that is more about beauty and aestheticism rather than promoting the same ideas and image that people usually associate with sex workers.

***Dissolution of Victory* by Daria Roman and Richard Essink** for Miglė Gerčaitė's *Remembering Soviet Lithuania in Grūtas Park.* Under Soviet rule, Soviet realist sculptures served to underline the clear superiority of the communist political system. Now, in the post-communist era, guided by sites of remembrance and reevaluation, like Grüntas Park, these statues lose that clarity. It dissolves into a more nuanced evaluation of their meaning. This is mirrored in the visual dissolution of the victory statue.

Contents

Sciences

A Quantitative Analysis of Diel Variation in Macroinvertebrate Assemblages in the Coastal Waters of Lipsi Island, Greece

The Importance of Day and Night Data for Marine Invertebrate Biodiversity Assessments

Gesa Carolin Mueller

Supervisor
Dr. Alison Gilbert (AUC)
Reader
Dr. Michiel van Drunen (AUC)

Photographer: Stef Deuring

Abstract

Marine macroinvertebrates make up over 95% of all marine animal species, occupying important ecological niches in coastal marine ecosystems and forming the basis for a majority of ecosystem services. However, most macroinvertebrate species are 'out of sight, out of mind' and largely underrepresented in science, policy, and conservation. This is of particular concern for nocturnally active species and raises the question of whether biodiversity and status assessments underestimate species richness and abundance by neglecting diel variation in activity patterns. Therefore, this study aims to contribute to the macroinvertebrate baseline inventory in an ecologically important yet understudied region and to investigate diel variation by (1) reporting all mobile and sessile species, and (2) quantifying differences in recorded species richness and abundance between day and night. Primary data were collected in the coastal waters around Lipsi island, Greece, during winter and early spring (February to April 2021) using underwater visual census. A total of 94 macroinvertebrate species from eight phyla were recorded, with an increase in species richness of 30.14% during night surveys compared to day surveys. Diel variation in visible macroinvertebrate assemblages was confirmed by the statistical analysis, which revealed that mobile species were more than twice as abundant during the night. Lastly, a number of the recorded species were selected to review and illustrate their ecological importance, ecosystem services provided, and implications for policy and conservation. The findings of this study emphasize the need to conduct both day and night surveys to account for nocturnally active species and increase representation in status assessments and conservation efforts.

Keywords: *Marine macroinvertebrates, biodiversity, diel variation, Mediterranean Sea*

Suggested citation: Mueller, G. (2021). A quantitative analysis of diel variation in macroinvertebrate assemblages in the coastal waters of Lipsi island, Greece: The importance of day and night data for marine invertebrate biodiversity assessments. *AUC Undergraduate Journal of Liberal Arts and Sciences, 16*(1), 1-29.

Contents

List of Figures and Tables

Abbreviations

CBD	Convention on Biological Diversity
CoE	Council of Europe
DD	Data Deficient
EC	European Commission
EEA	European Environment Agency
EU	European Union
GES	Good Environmental Status
IPBES	Intergovernmental Science-Policy Platform on Biodiversity
IPCC	Intergovernmental Panel on Climate Change
IUCN	International Union for Conservation of Nature and Natural Resources
MA	Millennium Ecosystem Assessment
MAA	Mycosporine-like Amino Acids
MEGR	Greek Ministry of the Environment, Energy and Climate Change
MPA	Marine Protected Area
MSFD	Marine Strategy Framework Directive
NGO	Non-Governmental Organization
SAC	Special Areas of Conservation (designated by the EU Habitat Directive)
UN	United Nations
UVC	Underwater Visual Census

Definitions

Biodiversity or *biological diversity* is defined as "the variability among living organisms from all sources [...]" by the UN Convention on Biological Diversity (CBD 1992, p. 3). In this study, biodiversity is assessed in terms of species richness, Shannon diversity, and evenness of macroinvertebrates (Wilsey and Potvin 2000).

Biogenic habitats are habitats formed by benthic organisms adding to the structural complexity of an ecosystem (Dunham et al. 2018).

Diel variation refers to changes in environmental factors (e.g., light, temperature) and activity patterns of organisms within a 24-hour period. For the purpose of this study, the term describes changes in the observed presence of a species between day and night surveys.

Ecological niche refers to a set of environmental conditions under which a species can exist indefinitely (Hutchinson 1957).

Ecosystem engineers are species that modify or maintain habitats and modulate the availability of biotic and abiotic materials to other species, either directly or indirectly (Jones et al. 1994).

Ecosystem functions refer to the biological properties and processes of an ecosystem (Costanza et al. 1997).

Ecosystem services are services for humans provided by biological diversity (provisioning, regulating, cultural, and supporting). These services emphasize that "biodiversity and human well-being are inextricably linked" (MA 2005, p. iii).

Flagship species are referred to as charismatic species that may serve as an ambassador or symbol for conservation campaigns to stimulate action (Heywood 1995).

Good environmental status refers to Article 3(5) of the MSFD and describes a productive, healthy, and diverse environmental status of oceans and seas, which requires the sustainable use of marine waters. The GES is evaluated based on eleven qualitative descriptors (EU 2008).

Macroinvertebrates are (predominantly) bottom-dwelling animals without a backbone visible to the naked eye (Brusca and Brusca 2003).

Non-indigenous species are defined as species introduced into habitats outside of their native range (MEECC 2014).

Oligotrophic describes an environment that is characterized by low nutrient levels (Ignatiades 1998).

Primary production forms the base of marine food webs and can be defined as the conversion of inorganic matter into new organic matter by primary producers (autotrophs) through the process of photosynthesis or chemosynthesis (Webb 2019).

Secondary production refers to the formation of biomass by consumers (heterotrophs), driven by the transfer of energy in the form of organic materials between trophic levels (Benke and Huryn 2006).

Sublittoral zone refers to the marine area below the intertidal zone, extending to a depth of 200 m. The sublittoral zone is permanently covered by seawater but sunlight still reaches the sea bed (Lincoln Smith 1988).

Protected areas are geographically defined areas that are regulated and managed for specific conservation strategies and aims (CBD 1992).

> *"The truth is that we need invertebrates but they don't need us. If human beings were to disappear tomorrow, the world would go on with little change. [...] But if invertebrates were to disappear, I doubt that the human species could last more than a few months. Most of the fishes, amphibians, birds, and mammals would crash to extinction about the same time. [...]"*
> — Wilson 1987, p. 345

1 Introduction

Invertebrates inhabiting the marine biome were among the first metazoan life on Earth, dating back to the Ediacaran period 635 million years ago. Molecular fossils suggest that these early animals were sponges (Gold et al. 2016), and by the end of the Ediacaran period, invertebrates of the modern phyla Porifera, Cnidaria, and Mollusca were already well established in the oceans (Brusca et al. 2016). During the Cambrian explosion, roughly 539 million years ago, invertebrates continued to evolve into a highly diverse group of animals, still constituting the only multicellular life in the sea (Brusca et al. 2016). The vast diversity of invertebrate species was already documented by the Greek philosopher and scientist Aristotle based on his observations in the Aegean Sea over 25 centuries ago (Voultsiadou and Vafidis 2007; Voultsiadou et al. 2010). Today, marine invertebrates account for over 95% of all marine animal species (Collen et al. 2012). They are fundamentally important for marine ecosystems, supporting biodiversity and key ecological functions.

Humans rely on the regulating and provisioning services of marine invertebrates, often unknowingly (Prather et al. 2013). 'Out of sight, out of mind' may be one of the greatest threats to this diverse group of species, hindering conservation and monitoring efforts (Collier et al. 2016; Eisenhauer et al. 2019). There is a general lack of awareness and knowledge of the importance and ecological role of marine invertebrates in comparison to the charismatic megafauna (Collier et al. 2016). This is of particular concern for the wide range of nocturnally active species that remain concealed during the day as they are rarely seen by the public and missing from many scientific reports. Species are threatened by habitat loss and fragmentation (Coll et al. 2010), overexploitation (Tsikliras et al. 2013), and ocean acidification (IPCC 2019), yet their representation in status assessments, policy plans, and conservation strategies is insufficient (Collen et al. 2012; IUCN 2021).

The present study aims to investigate biodiversity and patterns of diel variation in marine macroinvertebrate assemblages around Lipsi island, Greece, during winter and early spring (February to April). Primary data on the day and night species richness and abundance of macroinvertebrates were collected using underwater visual census surveys to address the following research objectives:

1. Assess macroinvertebrate diversity[1] and species richness around Lipsi island and contribute to the baseline inventory of an understudied yet ecologically important region.
2. Quantitatively investigate diel variation in mobile macroinvertebrate species richness and abundance using statistical analysis and address the question of whether diurnal sampling efforts are sufficient for status assessment, conservation, and monitoring programs.
3. Qualitatively assess the implications of reported biodiversity and diel variation with respect to species' ecological importance, their provision of ecosystem services, and implications for policy and conservation.

It is hypothesized that species richness and abundance are underestimated when neglecting nocturnally active species. While spatial, seasonal, and lunar variation may also influence activity patterns of invertebrates, this study focuses on diel variation; the examination of other environmental factors is beyond the scope of this study.

The present study can be divided into four main sections following this introduction. The first section contextualizes the research by reviewing the existing literature on marine invertebrates, biodiversity assessments, and associated policies with a focus on the Mediterranean and Aegean Seas. Section two details the study area, the survey methodology employed to collect in situ primary data, and the data analysis performed. Section three outlines the findings of the study, first in an empirical manner (i.e., species and phyla observed), and then quantitatively by calculating biodiversity and applying statistical analyses. The fourth section addresses each research objective and situates findings in the context of ecology and conservation. The final section summarizes the results of the study, reflects on the research limitations, and considers further needs for research.

2 Research Context

2.1 Marine macroinvertebrates

The term 'invertebrate' refers to those species in the kingdom Animalia that lack a vertebral column (Brusca et al. 2016). Their size ranges from Loriciferans, as small as 0.085 mm, to the giant squid (Architeuthis dux), reaching a size of 18 m (Brusca et al. 2016). 'Macroinvertebrates' describe species visible to the naked eye and are the focus of the present study. Marine macroinvertebrates (hereafter 'macroinvertebrates') can be categorized by their location in

[1] Diversity, in this study, refers to alpha diversity, i.e., the diversity within a particular habitat or sample (Whittaker 1972).

the water column: Benthic organisms live on the sea bottom and can be further distinguished into sessile (attached to the substratum; e.g., barnacles), sedentary (unattached but largely inactive; e.g., crinoids), and errant (active; e.g., crabs) species, depending on their locomotor capabilities (Brusca et al. 2016). Pelagic organisms inhabit the water column, either drifting with the water movement (e.g., jellyfish) or actively swimming (e.g., octopus and squid). For the purpose of this study, pelagic and sedentary or errant benthic species will be referred to as 'mobile invertebrates.' Species that are attached to the substratum will be referred to as 'sessile invertebrates.'

2.1.1 Ecosystem functions and services

Marine invertebrates are fundamentally important to ecological processes and ecosystem services (Prather et al. 2013). As a highly diverse group of organisms, invertebrates occupy a correspondingly wide range of niches, thus facilitating the coexistence of many specialized species (Wilson 1987). Additionally, numerous species are considered ecosystem engineers, providing complex habitats and modulating the availability of biotic and abiotic materials to other species, either directly or indirectly (Jones et al. 1994). In the Mediterranean Sea—the second largest marine and coastal biodiversity hotspot (Gabrié et al. 2012)—benthic organisms such as sponges, which are highly abundant in the region, modify the seascape and form biogenic reefs, providing microhabitats for other species (Bell 2008). Besides having evolved as some of the earliest animals, marine sponges also carry out a significant number of functional roles. For example, sponges facilitate carbon flow between the benthic and pelagic environments (De Goeij et al. 2013) and increase both primary and secondary production (Wulff 2006). Likewise, other marine invertebrates such as echinoderms and bivalves increase nutrient quality through decomposition, making detritus easily consumable by microorganisms (Prather et al. 2013).

Countries rely on their 'natural capital' – ecological resources and environmental assets in the form of ecosystems, species, and habitats that benefit humans in various ways, both directly and indirectly (Collen et al. 2012). Coastal marine ecosystems, in particular, are characterized by high productivity, supporting the livelihood of local fisheries (OECD 2021), protecting shorelines (Rife 2014), and sequestering carbon (Beaumont et al. 2014). These ecosystem services can be categorized into supporting, provisioning, regulating, and cultural services (Table 1), as defined by the Millennium Ecosystem Assessment (MA 2005). Marine invertebrates form the basis for a majority of goods and services, yet they remain largely unseen in comparison to the more charismatic marine megafauna (Collier et al. 2016). According to an analysis by Costa Leal et al. (2012), more than 75% of marine natural products, including pharmaceuticals treating chronic diseases, have been sourced from the phyla Porifera and Cnidaria since the 1990s. Oysters, shrimp, squid, and other species are commercially fished and consumed, providing food and income for many countries (Collen et al. 2012). In 2018, Greek fisheries produced seafood (fish, molluscs, crustaceans) with a value of USD 158.8 million, creating employment for approximately 20.000 people (OECD 2021).

Table 1: Ecosystem services provided by marine invertebrates. Adapted from Prather et al. 2013.

Type of service	Ecosystem service	Examples	References
Supporting	*Decomposition* *Bioturbation* *Nutrient-cycling* *Habitat formation*	Detritus-feeders such as holothurians improve nutrientcycling and bioturbation; many sponges form biogenic reefs	Amon and Herndl 1991; Bell 2008; De Goeij et al. 2013; Alexander et al. 2014
Provisioning	*Natural products (food, pharmaceuticals)*	Numerous sponges and sea anemones provide bioactive compounds for pharmaceuticals; molluscs are a source of food and income	Hassan et al. 2004; Koutsoubas et al. 2007; Costa Leal et al. 2012; Stabili et al. 2015
Regulating	*Water quality* *Food web stability*	Suspension feeders such as sponges and bivalves improve the water quality through their filtering activity	Officer et al. 1982; Ostroumov 2005; Bell 2008
Cultural	*Recreation*	Coral reefs for recreational diving	Costanza et al. 1997; Voultsiadou et al. 2010

2.2 Diel variation

2.2.1 Circadian rhythms in the animal kingdom

The natural environment is a complex system influenced by the periodicity of the sun and the moon, resulting in circadian, seasonal, lunar, and tidal rhythms, with the daily light-dark cycle as one of the most important cycles in the physical environment (Rosbash and Hall 1989; Tessmar-Raible et al. 2011). Rosbash (2017) defined circadian rhythms as "biochemical, physiological, behavioural adaptations to external daily oscillations" that have evolved over millions of years to anticipate daily external change. Simplified, while the endogenous clock in animals (including humans) does not exactly resemble the 24-hour periodicity of the external environment (hence circadian rhythm), it is entrained by input signals such as light and temperature, also referred to as 'zeitgeber' (Fig. 1; ref; Rosbash 2017), thus establishing a rhythm that closely mimics the length of the day. Negative feedback loops within the endogenous oscillator, or pacemaker, control the circadian timekeeping through the transcription of a 'clock gene' and the translation of its RNA into protein (Hardin et al. 1990). If external cues change, organisms tend to align to the new rhythm as a result of the coupling between the endogenous and exogenous oscillatory systems (Roenneberg et al. 2003). The synchronicity of an organism with both its internal and external stimuli is crucial for survival in an environment with complex ecosystem dynamics (Rosbash 2017).

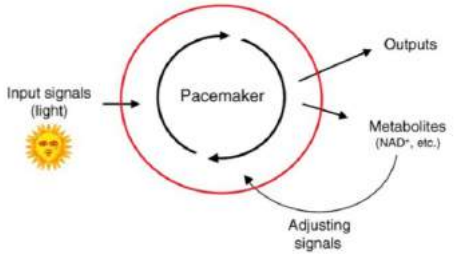

Figure 1: Graph visualizing the simplified concept of endogenous circadian clocks, influenced by internal and external stimuli (Rosbash 2017).

Research suggests that all marine species show patterns of rhythmicity generated by the endogenous oscillator within an organism and by external cycles such as the light-dark cycle (Aguzzi et al. 2012). The location of pacemakers in the nervous system of invertebrates is species-dependent, and multiple endogenous oscillators may exist within an organism (Aréchiga et al. 1993). For example, experiments found pacemakers located inside the retina of several molluscs and inside the brain (supraesophageal ganglion) of arthropods (Aréchiga et al. 1993). In numerous marine invertebrates such as flatworms and the sea star *Echinaster brasiliensis*, endogenous melatonin production has been identified as the primary mechanism regulating the circadian rhythm with a nocturnal peak (Itoh et al. 1999; Peres et al. 2014). However, these patterns are still poorly understood in the majority of marine invertebrates (Aguzzi et al. 2012).

2.2.2 Time as an ecological niche

Kronfeld-Schor and Dayan (2003) describe time as an ecological niche axis facilitating the coexistence of organisms. Ecological niches are a fundamental concept proposed by Hutchinson in 1957, describing a set of environmental conditions under which a species can exist indefinitely. However, the traditional niche concept is largely focused on the idea that biotic and abiotic factors influence population dynamics on a spatial scale (Holt 2009). Time as a niche with changing biotic and abiotic factors may lead to the 'temporal partitioning' of species, reducing resource competition and predator-prey interactions in communities, thus potentially explaining the evolutionary significance of endogenous rhythmicity (Schoener 1974). The ability to anticipate environmental changes such as light and darkness enables an organism to choose the most optimal time for a given behaviour or activity (Kronfeld-Schor and Dayan 2003).

Studies investigating circadian feeding and activity patterns in marine invertebrates found that many benthic and pelagic invertebrate taxa seek refuge during the day and only emerge at night (Blackmon and Valentine 2013; Brewin et al. 2016; Hinojosa et al. 2020). For example, nocturnal activity patterns were found in shrimps (Ory et al. 2014), urchins (Dee et al. 2012), holothurians, and echinoids (Hammond 1982). These studies suggest that the nocturnal specialization of species may have evolved as a strategy to reduce food web interactions, either as a predator-avoidance mechanism (Hammond 1982) or a strategy to reduce resource competition (Kronfeld-Schor and Dayan 2003), depending on the trophic group of a species. Across all studies, the nocturnal emergence of mobile marine invertebrate species was found to result in significant species turnover between day and night, affecting ecosystem dynamics and predator-prey interactions (e.g., Hinojosa et al. 2020; see Fig. 2). However, despite the significant differences between day and night species assemblages reported in the studies mentioned above, research on biodiversity conducted in the past decade has largely neglected diel changes (Ryu et al. 2012; Bumbeer et al. 2016), perhaps due to the increased logistical efforts associated with conducting night surveys. Excluding nocturnally active invertebrates from sampling procedures may result in the underestimation of biodiversity and thus compromise conservation efforts and the understanding of ecological processes.

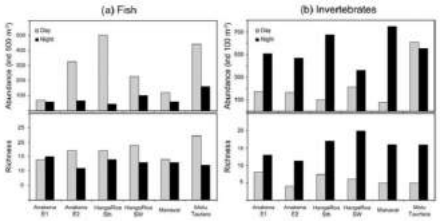

Figure 2: Comparison of day and night reef assemblages around the Easter Islands, supporting the hypothesis of temporal partitioning. Fish abundance was twice as high during the day, whereas invertebrate abundance was higher during the night (Hinojosa et al. 2020, p. 5).

2.3 Status assessment and policy

Despite their value for ecosystem functioning and services, marine invertebrates are underrepresented in monitoring programs and conservation targets (Collier et al. 2016; Eisenhauer et al. 2019). In the 1999 conference 'The Other 99%', it was already recognized that governments, society, and scientists have placed "far too much significance on the remaining 1%" (Reid 2000, p. 178), referring to flagship species that are considered 'charismatic' by the western culture and thus used as ambassadors for conservation campaigns (Heywood 1995; Barua 2011). While the use of flagship species, which are predominantly larger bodied mammals, may be effective in engaging the public, their use should improve rather than compromise the protection of the large diversity of invertebrates—the other 99%. This unbalance is also reflected in the European's funding instrument of the environment LIFE (EU 2013), as revealed in an analysis by Mammola et al. (2020): Between 1992 and 2018, financial investments were six times higher for vertebrates than for invertebrates, underlining the unequal distribution of conservation efforts. In global assessments by the IUCN Red List, 27% of invertebrates are "Data Deficient" (referred to as DD in Fig. 3; see also IUCN 2021). This is of particular concern given that marine invertebrates are among the most commonly overexploited species (MA 2005).

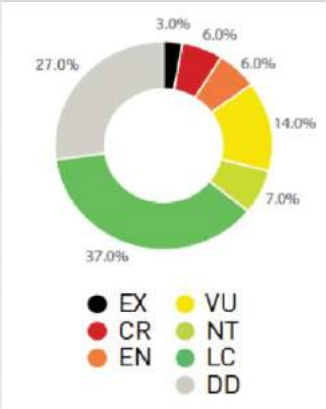

Figure 3: IUCN Red List assessment of marine invertebrates (Collen et al. 2012, p. 18).

Targets under the Convention on Biological Diversity (CBD) and the Marine Strategy Framework Directive (MSFD) have been established to address the global decline of biodiversity and achieve "Good Environmental Status" (GES) in EU waters (EU 2008). Among other targets, the GES aims to maintain biodiversity (Descriptor 1) and seafloor integrity (Descriptor 6), which requires ongoing assessments of ecosystems, habitats, and communities (EC 2010). However, marine invertebrates are omitted from reports under the MSFD (Palialexis et al. 2018) and the European Commission (EC 2018). The reports cover marine mammals, fish, birds, and reptiles but limit invertebrate assessments to cephalopods. The "invertebrate bottom fauna" targeted by Descriptor 6 (EU 2008 Annex III) is largely underrepresented in the reports (CoE 2006; EC 2018). According to Article 8(1) of the MSFD, all member states are requested to "implement coordinated monitoring programs for the ongoing assessment of the environmental status of their marine waters" (EU 2008). However, regional reports on Greece and the Aegean Sea are lacking data as Greece failed to submit information on monitoring programs (Dupont et al. 2014). The region is an important habitat for many species that are protected under the Habitats Directive (1992) and although the network of protected areas has been expanded in Greece (Gabrié et al. 2012), there is a lack of policy reinforcing biodiversity protection and management (Apostolopoulou and Pantis 2009). For example, management plans for the Special Areas of Conservation (SAC) Lipsi, Arki, Agathonisi, and Vrachonides in the Eastern Aegean Sea currently do not exist (EEA 2020). As a result, critically endangered invertebrate species such as *Tonna galea* and *Pinna nobilis* (CoE 2018) continue to be harvested for commercial use without legal consequences (Katsanevakis et al. 2011). Consequently, status assessments and ongoing monitoring programs for marine invertebrate biodiversity are needed to highlight the importance of increased conservation and legal protection.

2.4 Mediterranean Sea

The present study was conducted in the north-east of the Mediterranean Sea, which is the largest semi-enclosed basin on Earth, with a surface area of 2.5 million km2 and a coastline stretching over 46,000 km (Coll et al. 2010; Santinelli 2015). The Mediterranean Sea is the second-largest marine and coastal biodiversity hotspot and a global priority ecoregion (Gabrié et al. 2012) with significant environmental, economic, and cultural value for coastal communities (MA 2005; MEGR 2014). It is connected to the Atlantic Ocean through the Strait of Gibraltar, significantly affecting the oceanographic circulation in the region (Béthous 1979). As a concentration basin, the Mediterranean Sea is characterized by high evaporation exceeding precipitation, making it one of the saltiest seas on Earth (Lionello et al. 2006). In a broad sense, the circulation within the basin results from a pressure gradient from west to east, where cold, low-salinity water from the Atlantic is pushed across the Mediterranean basin (Coll et al. 2010; Schroeder et al. 2012). Temperature and salinity increase as the water flows eastwards towards the Levantine Sea, where the dense, saline water sinks and circulates back towards the West. Here, it exits the basin through the Strait of Gibraltar back into the Atlantic Ocean (Coll et al. 2010). Because of its enclosed nature and relatively small size, the Mediterranean Sea is sensitive to environmental and anthropogenic stressors such as climate change, overexploitation (Schroeder et al. 2012), and non-indigenous species (Zenetos et al. 2009). An analysis by Giorgi (2006) identified the Mediterranean as one of the most prominent climate change hotspots, threatening coastal marine ecosystems through sea level rise, increasing sea surface temperatures, and acidification (Coll et al. 2010).

2.4.1 The Aegean Sea

The Aegean Sea, located in the north-eastern Mediterranean between Greece and Turkey (Fig. 4), is considered a hotspot with particular ecological importance (Coll et al. 2010). Geographically, it is connected to the Black Sea through the Marmara Sea in the northeast, and bound by the Greek mainland to the west, the Anatolian peninsula to the east, and Eastern Macedonia to the north. The southern boundary is roughly marked by the deep trough south of Crete and Rhodes (Sini et al. 2017). The region covers approximately 191,000 km^2 and hosts a large number of islands, divided into seven island groups, including the Dodecanese and Cyclades. The climate in the region is characterized by cool and rainy winters from November to March, and dry, hot summers from May to September, with April and October exhibiting intermediate climate (Poulos et al. 1997). While species richness and marine diversity in the Mediterranean Sea largely show anorth-western to south-eastern gradient with decreasing productivity towards the east, the Aegean Sea constitutes an exception with high numbers of inver-

tebrates and marine mammals, including concentrations of threatened or endangered species (Coll et al. 2010). The high biodiversity in the otherwise oligotrophic Eastern Mediterranean may result from the Aegean deep water formation bringing nutrient-rich water to the surface (Sini et al. 2017). The region is also of historical interest and importance, as some of Aristotle's earliest work on marine biodiversity was based on his experience and observations on the island of Lesbos in the Aegean Sea (Voultsiadou and Vafidis 2007). Nevertheless, the Aegean Sea is still largely understudied and underrepresented in assessment reports and conservation strategies (Dupont et al. 2014).

The present study is an empirical effort to collect primary data on mobile and sessile macroinvertebrate species in the historically and ecologically important Aegean Sea around Lipsi island. The study thereby complements previous biodiversity assessments, which have largely focussed on a single phylum in a wider region rather than providing an overview of the macroinvertebrate diversity in a confined area. Moreover, this work is among the few biodiversity assessments to quantify diel variation in macroinvertebrate species richness and abundance by conducting day and night surveys, thus targeting nocturnally active species.

Figure 4: Map visualizing the geographic boundaries of the Aegean Sea, Greece (Sini et al. 2017, p.4).

3 Methodology

3.1 Survey sites

Primary data were collected at four sites around the island of Lipsi (37°18' N, 26°45' E) to investigate diel variation in marine macroinvertebrate assemblages. Lipsi island is located in a Natura 2000 SAC in the Eastern Aegean Sea, Greece (Fig. 5). The survey sites were selected based on previous protocols used by the Archipelagos Institute of Marine Conservation (Archipelagos Institute 2021), taking factors such as depth and wind exposure into consideration to include a wider variety of environmental conditions (Table 2). The selected survey sites are classified as "large shallow inlets and bays" (Natura 2000 code 1160) according to the

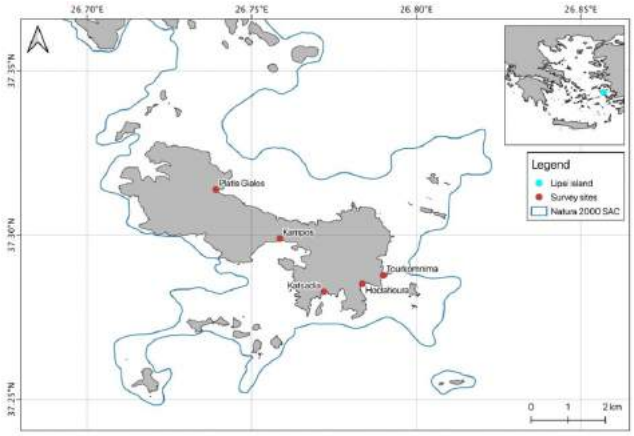

Figure 5: Lipsi island, in the Eastern Aegean Sea, Greece, and the location of the four survey sites. The blue line delineates the Natura 2000 SAC site Arki, Lipsi, Agathonisi and Vrachonisides (EEA 2020).

Habitats Directive (1992), with rocky substrates providing complex habitats for diverse benthic communities. Human presence was absent to low at all sites due to the reduced tourism in the winter season and COVID-19 related travel restrictions, limiting biases caused by anthropogenic disturbances (Usseglio 2015).

Table 2: Description of the four survey sites around Lipsi island.

Survey site	Location on Lipsi	Max. depth
Kampos	South, facing South	2.7 m
Hoklahoura	East, facing South-East	3.7 m
Katsadia	South, facing South	3.4 m
Platis Gialos	North, facing North East	1.7 m

3.2 Data collection

Each study site was replicated five times, resulting in a total of 20 day and night survey pairs conducted between February and April 2021. Survey pairs consisted of a day and night survey carried out at the same site within a 24-hour cycle to allow for accurate replication. Daytime surveys were carried out between 10:00h and 13:00h (UTC +2); nighttime surveys were carried out after sunset (between 18:00h and 21:00h; UTC + 2). Data were collected using standardized, non-invasive underwater visual census (UVC) methods (Brock 1954; Hinojosa et al. 2020). UVC techniques are the most widely used method in ecological surveys conducted in shallow and clear water as they are suitable for a wide range of species and habitats (Cheal and Thompson 1997; Edgar et al. 1997). Additionally, the method can be employed in marine protected areas (MPAs) due to its non-destructive nature (Garcia Charton et al 2000), making it appropriate for the data collection in the Natura 2000 SAC around Lipsi island. For safety reasons and to reduce inaccuracies caused by low visibility (Usseglio 2015), surveys were only conducted when weather conditions were favorable ($\leq$ Bft 4).

Two 50 m belt transects with a fixed width of 5 m were set up on rocky substrate in the sublittoral zone of each survey site (Fig. 6), resulting in a total census area of 225 m^2 at each site with a depth range of 1.0 to 3.7 m. Start- and endpoints of each transect were defined prior to the first survey to ensure consistent GPS coordinates. Environmental parameters (time, sea state, depth, and temperature) were recorded at the beginning of each survey using a Mares Smart Air dive computer. Surveys were carried out by two freedivers trained in the identification of marine invertebrates to minimize intra-observer biases (Bernard et al. 2013). The freedivers swam parallel to each other along both sides of the transect line, recording all visible macroinvertebrates ($\leq$ 1 cm in size) using a GoPro Hero 7 with a minimum resolution of 2.7k. The videos allowed for subsequent analysis of mobile and sessile species to increase the accuracy of species identification and counting. During night-time surveys, freedivers were equipped with a diving LED flashlight (165 lm), only recording mobile species as sessile species assemblages are not expected to vary between day and night.

Figure 6: Location of the 50 m transect lines at each survey site.

Surveys were followed by subsequent analysis of the recorded videos to identify and count observed species. Individuals were identified to species or, in some cases, genus level if identification to species was not possible. The identification was based on the regional macroinvertebrate ID guide compiled by the Archipelagos Institute for internal use (Archipelagos 2021). Additional species observed during the data collection that were not yet recorded in the ID guide were identified based on peer-reviewed literature (e.g., Chintiroglou et al. 2005; Voultsiadou 2005). Unidentified species were excluded from the dataset. Taxon details were retrieved from the World Register of Marine Species (WoRMS 2021).

3.3 Data analysis

Primary data were treated as incidence data, where *N* is the total numerical abundance (individual counts) across all surveys and $\overline{N}$ is the mean abundance (Seymour 2003). Data from the different study sites were combined as spatial variance was excluded from the analysis. For the principal analysis of diel variation, only mobile species were included to avoid biased results, as the presence and abundance of sessile species does not change between day and night. For the estimation of marine invertebrate diversity, both mobile and sessile species were included to represent the macroinvertebrate community inhabiting the coastal waters of Lipsi island.

3.3.1 Biodiversity estimation

Invertebrate assemblages were assessed in terms of diversity, species richness, and evenness, which are among the most widely used metrics of biodiversity in ecological studies (Thukral 2017). The diversity of mobile and sessile invertebrate assemblages was calculated using the Shannon Diversity Index *H*, defined as

$$H = -\sum_{i=1}^{s} p_i \ln p_i \tag{1}$$

where *s* is the total number of species (i.e., species richness) and p_i describes the proportion of the total sample to species *i* community (Shannon 1948). The index is a parameter commonly used in ecology to express the diversity of a given community (Seymour 2003). Evenness *E* describes the relative abundance of the species recorded in a given area (Wilsey and Potvin 2000) expressed as

$$E = \frac{H}{H_{max}} \tag{2}$$

where H_{max} is the maximum diversity possible (based on equation 1), given the number of species present.

3.3.2 Diel variation

Statistical analysis was performed using R software 4.0.2 (R Core Team 2020). Data were tested for normality using the Shapiro-Wilk Normality Test (Royston 1982). The test revealed that the dataset did not meet the assumptions of normality ($p < 0.001$); hence, the two-sided Wilcoxon Rank Sum Test (Neuhäuser 2011), a non-parametric alternative to the t-test, was performed to test for statistically significant differences in the presence and abundance of mobile macroinvertebrates between day and night. A pairwise multiple comparison test using the Wilcoxon Rank Sum Test was performed to examine which species exhibited significant differences in diel activity patterns. Multiple comparison tests tend to be prone to statistical errors; hence, P-values were examined to determine the risk of Type I error before drawing conclusions (Sato 1996).

In addition to the Wilcoxon analysis, a Hutcheson t-test was performed to compare day and night biodiversity for mobile species, as calculated using the Shannon Diversity Index. The Hutcheson t-test is given by

$$t = \frac{H_a - H_b}{\sqrt{s^2_{H_a} + s^2_{H_b}}} \quad (3)$$

where *H* represents the Shannon Diversity Index of two samples and *s* refers to the variance of each sample (Hutcheson 1970). The variables *a* and *b* correspond to day and night.

4 Results

4.1 Species richness and diversity

The present study collected primary data to contribute to the biodiversity inventory of marine macroinvertebrates around Lipsi island. A total of 94 species from eight different phyla (Fig. 7) were identified, of which 34 can be classified as sessile and 60 as mobile species (Fig. 8). While the observed species richness was higher in mobile species, sessile species were recorded in higher overall abundance ($\overline{N}_{sessile} = 52.8\%, \overline{N}_{mobile} = 47.2\%$; see Fig. 8). To quantify the overall diversity of the macroinvertebrate community found across all sites, the Shannon Diversity Index was calculated, revealing a value of $H = 3.01$. The overall species evenness was $E = 0.66$.

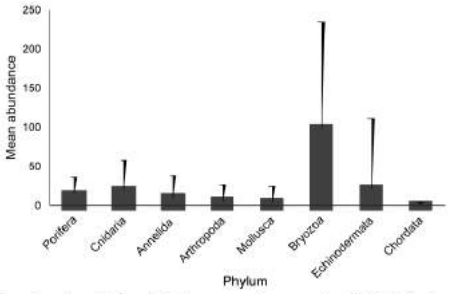

Figure 7: Mean abundance (count) of each phylum across all surveys ($n = 40$). Data is given as means ± SD.

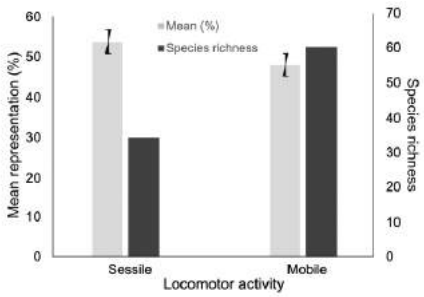

Figure 8: Representation (%) of locomotor activity across all surveys ($n = 40$), presented as mean ± SE (primary y-axis). Species richness across mobile and sessile species is displayed on the secondary y-axis (number count).

The most abundantly observed phyla ($\geq$ 10% of observed species) were Bryozoa, Echinodermata, and Cnidaria (Fig. 7). The phyla Mollusca, Porifera, and Echinodermata had the highest species richness of 37, 15, and 13 species, respectively. This finding is consistent with overall biodiversity estimations of the Mediterranean Sea (Coll et al. 2010). Bryozoa comprised almost half (47%) of the average species count. However, the highly abundant bryozoan species *Schizobranchiella sanguinea* ($\overline{N}$ = 62.65 ± 79.14; see Appendix II, Fig. 1AI) and *Reptadeonella violacea* ($\overline{N}$ = 210.8 ± 169.86; see Appendix II, Fig. 1AF) are colonial invertebrates, and spatial coverage assessed through quadrat sampling is a more accurate sampling method to estimate their abundance (Sullivan and Chiappone 1992). Therefore, the abundance of these species may have been overestimated in the present study. Since all colonial species recorded in this study are sessile macroinvertebrates excluded from the statistical analysis of day and night differences, the inaccuracy in estimating their abundance only affects the biodiversity assessment and not the quantification of diel variation. The abundance of the phylum Echinodermata (12.4%) can largely be attributed to the sea urchin *Paracentrotus lividus* (see Appendix II, Fig. 1AC), a species that was recorded in high abundance during both day and night surveys ($\overline{N}$ = 248.5 ± 217.7). Similarly, the solitary coral *B. europaea* (see Appendix II, Fig. 1C) made up the majority of the phylum Cnidaria ($\overline{N}$ = 58.7 ± 44.82).

4.2 Diel variation

The observed marine invertebrate abundance increased during night surveys, as visualized in Figure 9. Mobile invertebrates were more than twice as abundant during nocturnal surveys ($\overline{N}$ = 755.65 ± 335.17 indiv. per survey) compared to diurnal surveys ($\overline{N}$ = 340.8 ± 210.91 indiv. per survey; Fig. 10). This finding was confirmed by the statistical analysis, which revealed that day and night differences were statistically significant (Wilcoxon, $p < 0.001$).

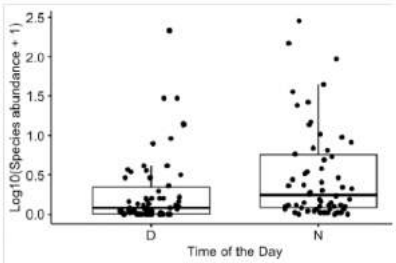

Figure 9: Boxplot visualizing the aggregated mean data of all mobile species observed on a log-scale ($n = 20$).

Statistically significant diel variation was found for 21 mobile species recorded during all day and night survey pairs across five taxonomic groups ($p < 0.05$; Table 3). The risk of a type I error was considered low due to consistently low p-values. With the exception of the nudibranch *Elysia timida* (see Appendix II, Fig. 1N), which was observed in higher

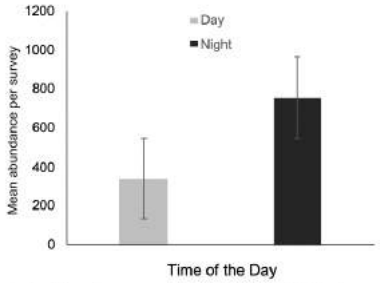

Figure 10: Bar chart visualizing the average number of mobile individuals observed during a day and night survey. Data is given as means ± SE (n = 20).

abundance during daytime sampling, all species were significantly more abundant during nocturnal survey efforts. The overall species richness increased by 30.14% between day and night, with a total of 44 mobile species observed during the day and 57 species observed at night. Sixteen species were observed exclusively at night and three exclusively during the day (Table 4). However, for seven of the nocturnally observed and all three of the diurnally detected species, observations were insufficient to draw conclusions about their temporal activity patterns ($\leq$ 5 observations across all surveys). However, for seven of the nocturnally observed and all three of the diurnally detected species, observations were insufficient to draw conclusions about their temporal activity patterns ($\leq$ 5 observations across all surveys).

Table 3: Wilcoxon Rank Sum output performed for each species on the aggregated dataset (mobile species only, presented in alphabetical order).

Species	*p*-value	Mean abundance and standard deviation		
		Day	*Night*	
Actinia equina	< 0.0001	0 (± 0)	2.8 (± 2.93)	****
Astropecten spinulosus	0.009	0.1 (± 0.31)	1.3 (± 2.23)	**
Bonellia viridis	< 0.0001	0 (± 0)	43.45 (± 38.85)	****
Calcinus tubularis	< 0.0001	0.3 (± 0.73)	5.5 (± 7.31)	***
Callistoctopus macropus	0.003	0 (± 0)	0.4 (± 0.5)	**
Cerithium spp.	0.02	13.2 (± 7.04)	23 (± 15.24)	*
Clibanarius erythropus	< 0.0001	0 (± 0)	2.45 (± 2.8)	****
Conus ventricosus	0.009	3.15 (± 3.25)	5.9 (± 4.27)	**
Dardanus arrosor	0.02	0 (± 0)	0.8 (± 1.61)	*
Elysia timida	< 0.0001	12.5 (± 10.04)	1.2 (± 1.58)	****
Eupolymnia nebulosa	< 0.0001	0.6 (± 1.27)	9.4 (± 9.29)	****
Hermodice carunculata	< 0.0001	2.45 (± 3.56)	92.1 (± 55.18)	****
Holothuria forskali	< 0.0001	1.9 (± 2.43)	34.6 (± 38.63)	****
Holothuria tubulosa	< 0.0001	0.1 (± 0.31)	4.8 (± 7.11)	****
Marthasterias glacialis	0.002	0.1 (± 0.31)	0.9 (± 1.02)	**
Ophiothrix fragilis	< 0.0001	0.6 (± 0.94)	8.5 (± 7.68)	****
Ophioderma longicauda	< 0.0001	0 (± 0)	4.35 (± 7.34)	****
Pachygraspus marmoratus	0.04	0 (± 0)	0.3 (± 0.73)	*
Pagurus anachoretus	< 0.0001	0.5 (± 0.76)	13.75 (± 11.68)	****
Pyrene scripta	< 0.0001	28.55 (± 21.94)	145.95 (± 82.16)	****
Rhyssoplax olivacea	0.005	0.6 (± 1.47)	3.9 (± 5.19)	**

Significance levels are indicated by: ****p < 0.0001; ***p < 0.001; **p < 0.01; *p < 0.05.

The Hutcheson t-test revealed statistically significant differences in day and night Shannon Diversity for mobile species, t(579) = 5.723, $p < 0.001$ (Table 4; Fig. 11), with observed mobile species assemblages showing higher biodiversity during nocturnal surveys ($H_{night} = 2.18$) compared to diurnal surveys ($H_{day} = 1.58$).

Table 4: Day and night comparison of the Shannon Diversity Index for mobile species, using the Hutcheson ttest (n = 20).

	Mean abundance	Richness *S*	Shannon Diversity *H*	Variance S_{2H}	*t*	*df*	Critical value	*p*-value
Day	340.8	44	1.58	0.008	5.723	579	1.96	< 0.0001
Night	755.65	57	2.18	0.002				

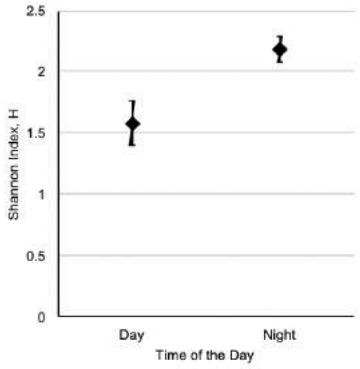

Figure 11: Shannon Diversity Index H comparing day and night for mobile species (n = 20).

5 Discussion

5.1 Marine macroinvertebrates diversity

Diversity and ecology of marine macroinvertebrates remain largely understudied in comparison to most vertebrate groups (Collier et al. 2016; Eisenhauer et al. 2019), suggesting the need for baseline inventories and monitoring programs. This study provides a comprehensive list of marine macroinvertebrates focusing on an island in the Natura 2000 SAC Arki, Lipsi, Agathonisi, and Vrachonisides. Biodiversity was expressed in terms of species richness and Shannon diversity. Across all 20 pairs of day and night surveys, a total of 94 macroinvertebrate species were observed around Lipsi island, with an overall Shannon diversity of $H = 3.01$. In ecological studies, the Shannon index generally lies between 1.5 and 3.5, with a high value representing greater diversity and evenness in a community (Wilsey and Potvin 2000). In comparison to other ecological studies quantifying invertebrate diversity using the Shannon index (e.g., Van Nguyen et al.; Wang et al. 2020), the diversity recorded in this study can be interpreted as medium to high (Chainho et al. 2007), reaffirming the ecological importance of coastal habitats in the region (Coll et al. 2010). The overall species evenness was $E = 0.66$, with only a few species recorded in particularly high abundances relative to the average species count (e.g., *P. lividus* and *R. violacea* with mean abundance counts of > 200 per survey), and numerous rare species (observed less than five times across all surveys, e.g., *Flabellina affinis* and *Thuridilla hopei*; see Appendix I, Table 1 and 2).

Previous biodiversity studies in the Aegean Sea focussed on single phyla rather than reporting the overall macroinvertebrate diversity in a confined area. For example, Voultsiadou (2005) and Zenetos et al. (2005) published reports on the diversity of sponges and molluscs in the Aegean Sea, which increased taxonomic knowledge and accommodated a wider range of natural variability but failed to provide information on overall invertebrate community structures and activity patterns. In comparison, the present study reports all phyla observed on rocky substrata in the sublittoral zone, however, it was limited to four study sites and thus possibly underestimates species richness.

5.2 Diel variation

With respect to the second objective of this report, the quantitative analysis of diel variation in mobile macroinvertebrate species richness and abundance, the present study showed significant differences in macroinvertebrate assemblages between day and night using statistical analyses. The mean abundance of macroinvertebrates was more than twice as high during nocturnal surveys compared to diurnal surveys. Notably, twenty mobile species were significantly more abundant at night, while only one species was reported in higher abundance during the day. Overall species richness increased by 30.14%, with 16 species exclusively observed at night. Similar trends were reported in recent studies conducted in the South Atlantic and the south-eastern Pacific: Brewin et al. (2016) found that invertebrate density increased by an average of 53.5% between diurnal and nocturnal surveying, and Hinojosa et al. (2020) similarly reported a three-fold increase in invertebrate abundance and richness at night. Both studies followed similar sampling procedures to the present study, using underwater visual census to record all visible species. In contrast, Blackmon and Valentine (2013) used day and night plankton tows in seagrass habitats to assess benthic macroinvertebrate emergence at night. While their study confirmed the previously discussed results reporting a 20-fold increase in macroinvertebrate density, towing techniques limit the data collection to species in the water column and neglect species on the benthos. Therefore, they are unsuitable for biodiversity assessments of benthic and pelagic species, particularly for those with rocky substrate habitats.

Previous studies included numerous environmental factors that were excluded from the scope of the present study, including seasonal variation, trophic dynamics, and lunar phases. For example, Brewin et al. (2016) collected data over a period of nine months, reporting strong seasonal patterns in addition to diel variation. Additionally, Brewin et al. (2016) and Hinojosa et al. (2020) analyzed both fish and invertebrate assemblages, revealing significant species turnover between day and night (see Fig. 2 in section 2.2.2) and underlining the importance to gain a better understanding of food web interactions and predator-prey dynamics. Lastly, Blackmon and Valentine (2013) controlled for the influence of moon phases, which are suggested to influence species assemblages significantly (Fingerman 1957). These findings encourage further in-depth investigations of macroinvertebrates and ecosystem dynamics around Lipsi island.

The results presented in this paper suggest that diel variation is a significant factor that should be accounted for in biodiversity assessments. Nevertheless, most studies collecting baseline data on marine invertebrate diversity are largely based on diurnal sampling efforts (e.g., Ryu et al. 2012; Bumbeer et al. 2016), neglecting species that only emerge at night. While statistical methods have been developed to account for "downward-biased" abundance estimations (Magurran and McGill 2011), undetected, nocturnally active species cannot be monitored and will be neglected in conservation and status assessments. Hence, the present study confirms the initial hypothesis that both day and night data collection is required to account for diel variation in activity patterns of macroinvertebrates.

5.2.1 Beyond biodiversity – relevance of selected species

A wide variety of mobile and sessile species were recorded, ranging from small, benthic species such as the highly abundant gastropod *Pyrene scripta* (see Appendix II, Fig. 1AE) or the nudibranch *Elysia timida* (approximately 1.5 cm in size) to larger, pelagic species such as the octopuses *Octopus vulgaris* (see Appendix II, Fig. 1AB) and *Callistocopus macropus* (reaching up to 150 cm; see Appendix II, Fig. 1F). Many of the observed species had already been recorded and classified by Aristotle, such as sponges of the genus *Sarcotragus spp* (see Appendix II, Fig. 1AH), the sea anemone *Actinia equina* (see Appendix II, Fig. 1A), the fireworm *Hermodice caranculata* (see Appendix II, Fig. 1Q), the highly abundant sea urchin *Paracentrotus lividus*, and numerous Molluscs (Voultsiadou and Vafidis 2007). The species discussed in greater detail below were selected to address the third objective of this research, highlighting the ecological importance, ecosystem services provided, and/or the relevance of marine invertebrates for policy and conservation.

5.2.2 Ecological functions and ecosystem services

While it would likely be possible to associate ecological functions and ecosystem services with a majority of the species encountered, this section only deals with a small selection of commonly observed species to exemplify the importance of the macroinvertebrates recorded around Lipsi island. The phylum Porifera was highly abundant across all survey sites (see Appendix I, Table 1) and fulfils critical ecological functions (see section 2.1.1). Sponges such as *Ircinia variabilis* (see Appendix II, Fig. 1S) increase the secondary production in a habitat as they are consumed by gastropod sea slugs and other species (Wulff 2006). Similarly, the detritus-

feeding sea cucumbers *Holothuria forskali* and *H. tubulosa* (see Appendix II, Fig. 1T and Fig. 1U) are important components of the benthic ecosystem as they stabilize bacterial communities in the sediment and contribute to nutrient recycling (Amon and Herndl 1991). The large keratose sponge *Sarcotragus spinosolus* (see Appendix II, Fig. 1AH) modulates the seascape and improves the water quality through its filtering activity. A study by Trani et al. (2021) found that *S. spinosolus* removed 99.72% of *Vibrio spp.*, a pathogen that is highly prevalent in aquaculture, from the water. Water filtration as a regulating ecosystem service is also performed by other suspension feeders recorded during this study, including bivalves and the highly abundant bryozoans *Reptadeonella violacea* and *Schizobranchiella sanguinea* (Officer et al. 1982). As reported in section 4.1, the phylum Bryozoa made up a significant part of the overall abundance counts (see Fig. 7), however, it also illustrates the underrepresentation of cryptic species in science and conservation (e.g., Pagès Escolà 2019). Provisioning ecosystem services (see Table 1 in section 2.1.1) can be illustrated by species including the sponge *Hamigera hamigera* (see Appendix II, Fig. 1P) and the sea anemone *Actinia equina*, which provide biologically active compounds used for pharmaceuticals (Hassan et al. 2004; Stabili et al. 2015). In Greece and other Mediterranean countries, the value of marine macroinvertebrates as 'natural capital' becomes evident in coastal communities that rely on the exploitation of molluscs such as *Octopus vulgaris* or the gastropod *Hexaplex trunculus* (see Appendix II, Fig. 1R) as a source of food and income (Koutsoubas et al. 2007).

5.2.3 Time as an ecological niche

Literature dealing with the 21 mobile species that showed statistically significant differences between day and night (see Table 3 in section 4.2) provides biological and ecological clarification for the observed activity and emergence patterns (e.g., Hammond 1982; Meisel et al. 2006). The only species that was significantly more abundant during the day was the gastropod *Elysia timida*, a sacoglossan (i.e., solar-powered) sea slug that lives in a kleptoplastic relationship with algal symbionts, incorporating chloroplasts from algal prey directly into its cell cytosol to utilize their photosynthetic function (Pelletreau et al. 2011). Hence, predominantly diurnal activity is required for the species to obtain energy (Giménez-Casalduero et al. 2011).

From the perspective of time as an ecological niche (Kronfeld-Schor and Dayan 2003), the coexistence of the two octopuses observed during this study, *Callistoctopus macropus* and *Octopus vulgaris*, illustrates how temporal partitioning can reduce resource competition. Both species are predators with overlapping habitats (Meisel et al. 2006) and are reported to be predominantly nocturnally active in the Mediterranean Sea (Boletzky et al. 2001; Brown et al. 2006). However, in this study, only *C. macropus* exhibited exclusively nocturnal activity patterns, whereas *O. vulgaris* was active both during the day and night. This finding corresponds with the results from a laboratory experiment by Meisel et al. (2006), who compared the activity patterns of both species and found that *O. vulgaris* changed from nocturnal to diurnal rhythmicity within three days in the laboratory setting, possibly due to entrainment of the endogenous clock to human activity (e.g. feeding times). In contrast, *C. macropus* did not change activity patterns in the laboratory setting and remained strictly nocturnal. This observation suggests that *O. vulgaris* may be adaptive to external cues such as light, providing a possible explanation for the diurnal foraging behavior observed around Lipsi island. Temporal partitioning in slow-moving species can also function as a predator-avoidance mechanism, as suggested in a study on holothurians by Hammond (1982). Consistent with findings by Hammond (1982), the abundance of holothurians reported in the present study (*Holothuria forskali* and *H. tubulosa*) increased drastically after sunset (see Table 3 in section 4.2), and individuals observed during daytime were usually hidden in caves or under rocks. However, Wheeling et al. (2007) proposed an alternative explanation for diel activity patterns in holothurians, namely the avoidance of UV radiation in warm regions with clear water. Mycosporine-like amino acid (MAA), a biochemical defense blocking harmful wavelengths, was found in very low concentrations in *Holothuria spp.*, suggesting that the species observed around Lipsi might be sensitive to UV radiation. In comparison, species in other studies (e.g., *Pearsonothuria graeffei*), which displayed higher activity rates during the day and remained exposed at all times, had among the highest concentration of MAA (Shick et al. 1992). Additionally, some of the main predators of holothurians, such as sea stars (Francour 1997), display predominantly nocturnal activity patterns (for example, *Astropecten spinulosus* and *Marthasterias glacialis*, see Table 3 in section 4.2; see also Appendix II, Fig. 1B and Fig. 1W), suggesting that predator-avoidance may not be the primary reason for the nocturnal emergence of the holothurians observed in this study.

The examples illustrated above highlight that circadian rhythms and the underlying reasons for diel variation in marine invertebrates are still poorly understood, and further research is needed to better understand ecosystem dynamics.

5.2.4 Status assessment and policy

The Aegean Sea was identified as a biodiversity hotspot with high concentrations of endangered or vulnerable species (Coll et al. 2010). Of the species recorded across all day and night survey pairs around Lipsi island, five species are protected or strictly protected under the Bern Convention (Table 4; CoE 1979 Annex II and III) and/or the European Habitats Directive (92/43/EEC). However, publications report that species such as *Lithophaga lithophaga* and *Tonna galea* (see Appendix II, Fig. 1V and Fig. 1AJ) are still illegally exploited and served in local restaurants in the region

(Katsanevakis et al. 2011), highlighting the need for stricter enforcement of policies.

Two non-indigenous species were recorded, *Diadema setosum* and *Chama pacifica* (see Appendix II, Fig. 1M and Fig. 1H), with the long-spined sea urchin *D. setosum* recorded at all survey sites. Both species are considered invasive in the region (Zenetos et al. 2010; Bronstein et al. 2017). While not recorded during this study, the invasive sea cucumber *Synaptula reciprocans* (Antoniadou and Vafidis 2009) was observed outside of the transect lines as well as during the previous summer (Archipelagos 2021, pers. comm.), emphasizing that the species recorded within the sampling area only constitute a subset of the total macroinvertebrate diversity around Lipsi island.

Table 5: Protected and non-indigenous or invasive species recorded during day and night surveys around Lipsi island.

Protected species (Council of Europe 1979, Annex II and III; European Habitats Directive (92/43/EEC)	***Non-indigenous and/or invasive species***
Strictly protected	*Invasive*
Charonia tritonis	Chama pacifica (Zenetos et al. 2010)
Lithophaga lithophaga	Diadema setosum (Bronstein et al. 2017)
Patella spp.	
Tonna galea	
Protected	
Paracentrotus lividus	

Coll et al. (2010) found that introductions of the phyla Mollusca (33%) and Arthropoda (18%) are most frequently reported in the Mediterranean Sea. However, the authors noted that overall data on non-indigenous invertebrates are likely downward biased as information on many smaller species is "entirely absent" (p. 9), suggesting that the data deficiency on marine invertebrates impairs the monitoring of ecosystem health with regard to ecological invasions. Similarly, the overall status assessment of the species recorded was compromised by inadequate treatment of marine invertebrates in policy reports (Dupont et al. 2014) and the IUCN Red List (Cardoso et al. 2011; IUCN 2021). While the IUCN Red List is among the most extensive resources on the population status of species, the majority of macroinvertebrates observed in this study are not present in the database, notably all species of the phyla Porifera and Bryozoa, and many species of the phyla Annelida, Mollusca, and Echinodermata. This exemplifies the previously discussed underrepresentation of marine invertebrates in policy frameworks and status assessments (see section 2.3) and underlines that data gaps and taxonomic biases ultimately must be addressed in EU biodiversity targets.

6 Conclusions

The present study was among the first to collect day and night winter data on marine macroinvertebrate assemblages around Lipsi island, Greece, using UVC transect surveys. The main aim of the study was to investigate biodiversity and diel variation in macroinvertebrates, posing the question of whether biodiversity and status assessments underestimate species richness and abundance when nocturnally active species are neglected. The aim was addressed by (1) assessing diversity and species richness of both sessile and mobile species, and (2) quantifying differences in species richness and abundance between day and night. Lastly, relevant species were discussed from the perspective of their ecological importance, the provision of ecosystem services, and conservation and policy implications.

Ninety-four species from eight different phyla were identified during the data collection between February and April 2021, providing an overview of the macroinvertebrate diversity on rocky sublittoral substrata in the Aegean Sea. Furthermore, the quantitative assessment of diel variation provided strong evidence to support the hypothesis that species richness and abundance are underestimated when neglecting nocturnally active species. Visible mobile macroinvertebrates were more than twice as abundant during nighttime compared to daytime data collection. Twenty mobile species showed statistically significant increases in abundance at night; in contrast, only one species was significantly more abundant during the day. The comparison of species richness and diversity across all sites and surveys showed a nocturnal increase in species richness of 30.14%, demonstrating that diurnal sampling efforts need to be augmented by nocturnal sampling for accurate estimation of macroinvertebrate biodiversity.

While the present study is unable to encompass a larger geographical area, it contributes to the biodiversity inventory of macroinvertebrates around Lipsi island. Moreover, the reported findings on diel variation and nocturnal activity patterns are relevant for conservationists and policy-makers as they contribute to the knowledge on a largely understudied group of animals and emphasize the need to re-evaluate sampling procedures and reduce research biases in future status assessment and monitoring programs.

6.1 Limitations

The findings of this report are subject to a number of limitations concerning both the data collection and analysis. Firstly, observational studies are considered "downward-biased estimators" of the total population size in an area (Seymour 2003) because they commonly underestimate species richness and abundance. In the present study, the data collected were limited to visible species within the survey area; hence, the 'false absence' of species (i.e., species are present but remained undetected) was a potential source of underestimation. There are statistical methods to correct for biased estimations; however, this study solely reported and quantified observations and did not correct for potential underestimation due to time limitations.

Secondly, in visual studies, observer bias due to limited diver efficiency cannot be excluded (Williams et al. 2006; Bernard et al. 2013). Bias derives from inter- and intra-observer biases, as well as factors such as experience level and/or physical constraints under cold conditions. Inter-observer bias, (i.e. variation among observers) was minimized by having the same two observers, similarly trained in marine macroinvertebrate identification, conduct all surveys. Intra-observer bias (e.g., missing cryptic species) was reduced by videoing each survey and performing subsequent analysis of the footage collected. Nevertheless, transparent or well-concealed species such as *Eupolymnia nebulosa* could have been easily missed. Combined with the difficulty of identifying certain species (particularly of the phyla Mollusca and Porifera), this study solely provides a snapshot of the overall biodiversity around Lipsi island.

Thirdly, bias resulting from behavioral responses by marine organisms (escape or attraction; see MacFarlane and King 2002; Usseglio 2015) cannot be excluded but was minimized by limiting site replications to one day-night survey per week. There was no indication that species reacted to the presence of the freedivers; hence, biases due to observer presence are unlikely. A number of octopuses were encountered several times at the same survey site and individuals interacted with the divers. However, time gaps between replications to minimize familiarity of the animals with divers, as well as studies on octopus behavior (Kuba et al. 2003), suggest that the species may simply be curious animals.

Lastly, a number of logistical restrictions must be acknowledged. Night-time surveys were conducted shortly after sunset; hence, data on nocturnal macroinvertebrate assemblages in this study do not consider activity and emergence peaks later at night. Furthermore, as mentioned in section 4.1, the sampling procedure for colonial organisms such as sponges and bryozoans can be improved by using quadrat sampling methods to estimate spatial coverage (Sullivan and Chiappone 1992). Finally, the study was limited to the shallow sublittoral zone at four survey sites, and the proposed fifth survey site, Tourkomnima, was excluded from the study due to its exposure to strong winds during the sampling period. While inter-site differences are likely minimal due to their geographic proximity, increased spatial coverage may reveal greater invertebrate diversity.

6.2 Future research

This study focussed on day and night data collection at four study sites during winter and early spring, leading to an initial assessment of biodiversity and diel variation around Lipsi island. Previous research on rhythmicity in marine invertebrates suggests that the presence/absence of species may also be influenced by other factors such as seasonal variation and lunar cycles. Hence, a natural progression of this work is the quantitative comparison of seasonal, annual, and lunar variation. Additionally, studies by Brewin et al. (2016) and Hinojosa et al. (2020) investigated diel variation in coastal ecosystems by recording both invertebrate and fish species. Including fish in the data collection could be expected to increase the understanding of predator-prey dynamics and temporal partitioning.

Acknowledgements

First and foremost, I would like to thank my friend, survey partner, and talented photographer, Janika Sander. Jan - I don't know what I would have done without you. Thank you for 50 + hours underwater filming invertebrates, for your help with my analysis, for good memories in very cold water, and for the beautiful photos you took during our surveys. I would also like to express my gratitude to the people from the Archipelagos Institute of Marine Conservation that supported me, especially Alice Malcolm-McKay and Laura Macrina. Alice, thank you for having been an amazing on-site supervisor and for introducing me to the world of marine invertebrates. My AUC supervisor, Dr. Alison Gilbert, had to keep up with many stories about changing topics, frostbites, and more. Thank you, Alison, for your patience and support. I am very grateful for the constant advice and feedback you provided. I would also like to thank my lecturers at AUC for the knowledge and skills they shared with me in the past three years, and my wonderful friends and family for supporting me throughout my degree. I am beyond grateful for the people I met during my time on Lipsi. This island has truly become a very special place for me, and it has been a beautiful and rewarding experience to collect my data in the bays around the island. I would like to thank my dear friend Lea for keeping me sane when I was stressed out about deadlines, and Ebba, who taught me some great wisdom about life ("If you feel too stressed to take a coffee break, then you should take an even longer break"). Finally, thank you, Kostas, for giving me a home on Dimitri's Farm while I was writing my capstone. You have created the most magical place on this island, and I cannot think of a better place to finish my thesis. The Sea holds a very special place in my heart, and I am eternally grateful for the glimpse into the underwater world I got to experience during my day and night surveys. Learning about the species I encountered has truly changed the way I look at the seascape and has taught me to pay attention to the small details. I hope, and I am determined, that my time working on Lipsi was just the beginning of a long journey committed to the conservation and protection of the ocean.

Appendix

I Recorded species

Table 1: Incidence list of all sessile species recorded across all day and night surveys.

Phylum Class	**Sessile species**	**Mean abundance**	**SD**
Annelida			
Polychaeta	*Bispira volutacornis* (Montagu 1804)	0.6	0.81
	Myxicola infundibulum (Montagu 1808)	0.325	0.66
	Protula tubularia (Montagu 1803)	22.15	17.88
	Sabella pavonina (Savigny 1822)	0.4	0.74
	Sabella spallanzanii (Gmelin 1791)	0.1	0.44
Arthropoda			
Maxillopoda	*Lepas anserifera* (Linnaeus 1767)	16.5	25.75
	Perforatus perforatus (Bruguière 1789)	35.4	40.36
Bryozoa			
Gymnolaemata	*Reptadeonella violacea* (Johnston 1847)	210.8	169.86
	Schizoporella dunkeri (Reuss 1848)	21.25	17.83
	Schizobrachiella sanguinea (Norman 1868)	62.65	79.14
Chordata			
Ascidiacea	*Ascidia mentula* (Müller 1776)	0.75	1.06
	Clavelina lepadiformis (Müller 1776)	0.2	0.61
	Halocynthia papillosa (Linnaeus 1767)	4.85	5.95
	Microcosmus polymorphus (Heller 1877)	0.8	0.99
	Microcosmus sabatieri (Roule 1885)	0.5	0.75
Cnidaria			
Anthozoa	*Balanophyllia europaea* (Risso 1862)	58.7	44.82
	Cerianthus membranaceus (Gmelin 1791)	0.075	0.27
Mollusca			
Bivalvia	*Ostrea edulis* (Linnaeus 1758)	3.23	3.56
Gastropoda	*Serpulorbis arenarius* (Linnaeus 1758)	1.63	2.41
Porifera			
Demospongiae	*Aplysina aerophoba* (Nardo 1833)	1.05	0.99
	Chondrilla nucula (Schmidt 1862)	0.35	0.48
	Chondrosia reniformis (Nardo 1847)	18.8	28.4
	Cliona celata (Grant 1826)	12.8	11.23
	Cliona schmidtii (Ridley 1881)	11.1	15.15
	Cliona viridis (Schmidt 1862)	15.95	15.45
	Crambe crambe (Schmidt 1862)	11.25	7.64
	Haliclona fulva (Topsent 1893)	12.95	9.66
	Hamigera hamigera (Schmidt 1862)	37.8	25.33
	Hemimycale columella (Bowerbank 1874)	0.75	0.78
	Ircinia variabilis (Schmidt 1862)	58.45	26.48
	Petrosia ficiformis (Poiret 1789)	0.65	1.29
	Phorbas topsenti (Vacelet & Pérez 2008)	3.3	4.31
	Sarcotragus spinulosus (Schmidt 1862)	21.95	17.88
	Spirastrella cunctatrix (Schmidt 1868)	16.25	13.89
Total 7	**34**		

Table 2: Incidence list of all mobile species recorded across all day and night surveys, including time of the day the species were observed. A minus (-) indicates that a given species was never observed across all 20 day or night surveys; a plus (+) indicates that a given species was observed at least once.

Phylum Class	**Mobile species**	**Mean abundance**	**SD**	**Presence**	
				Day	*Night*
Annelida					
Polychaeta	*Bonellia viridis* (Rolando 1822)	21.73	34.92	-	+
	Eupolymnia nebulosa (Montagu 1819)	5	7.92	+	+
	Harmothoe sp. (Kinberg 1856)	0.03	0.16	-	+
	Hermodice carunculata (Pallas 1766)	47.28	59.59	+	+
Arthropoda					
Malacostraca	*Calcinus tubularis* (Linnaeus 1767)	2.9	5.76	+	+
	Clibanarius erythropus (Latreille 1818)	1.23	2.31	-	+
	Dardanus arrosor (Herbst 1796)	0.4	1.19	-	+
	Gnathophyllum elegans (Risso 1816)	0.03	0.16	-	+
	Pachygrapsus marmoratus (Fabricius 1787)	0.15	0.533	-	+
	Pagurus anachoretus (Risso 1827)	7.13	10.57	+	+
	Palaemon serratus (Pennant 1777)	0.8	1.22	+	+
Cnidaria					
Anthozoa	*Actinia equina* (Linnaeus 1758)	1.4	2.49	-	+
Echinodermata					
Asteroidea	*Asterina gibbosa* (Pennant 1777)	0.03	0.16	-	+
	Astropecten spinulosus (Philippi 1837)	0.7	1.68	+	+
	Echinaster sepositus (Retzius 1783)	0.2	0.41	+	+
	Marthasterias glacialis (Linnaeus 1758)	0.5	0.85	+	+
Echinoidea					
	Arbacia lixula (Linnaeus 1758)	10.375	13.88	+	+
	Diadema setosum (Leske 1778)	1.78	2.54	+	+
	Paracentrotus lividus (Lamarck 1816)	248.05	214.7	+	+
	Sphaerechinus granularis (Lamarck 1816)	0.43	0.59	+	+
Holothuroidea	*Holothuria forskali* (Delle Chiaje 1823)	18.25	31.69	+	+
	Holothuria tubulosa (Gmelin 1791)	2.45	5.51	+	+
Ophiuroidea	*Ophiothrix fragilis* (Abildgaard, in Müller 1789)	4.55	6.72	+	+
	Ophioderma longicauda (Bruzelius 1805)	2.18	5.58	-	+
	Ophiocomina nigra (Abildgaard, in Müller 1789)	0.05	0.22	-	+

Table 2: Continued.

Phylum Class	**Mobile species**	**Mean abundance**	**SD**	**Presence**	
Mollusca					
Bivalvia	*Arca noae* (Linnaeus 1758)	0.35	0.58	+	+
	Barbatia barbata (Linnaeus 1758)	2.53	2.61	+	+
	Bolinus brandaris (Linnaeus 1758)	0.1	0.30	+	+
	Chama pacifica (Broderip 1835)	0.05	0.22	+	-
	Lithophaga lithophaga (Linnaeus 1758)	0.33	0.69	+	+
	Mimachlamys varia (Linnaeus 1758)	0.1	0.38	+	+
	Mytilus galloprovincialis (Lamarck 1819)	0.15	0.36	+	+
	Pecten jacobaeus (Linnaeus 1758)	0.08	0.267	+	+
	Pinctada imbricata radiata (Leach 1814)	0.33	0.66	+	+
	Pitar rudis (Poli 1795)	0.15	0.43	+	+
	Spondylus gaerderopus (Linnaeus 1758)	0.6	0.98	+	+
	Venus verrucosa (Linnaeus 1758)	0.83	1.19	+	+
Cephalopoda	*Octopus vulgaris* (Cuvier 1797)	0.33	0.62	+	+
	Callistoctopus macropus (Risso 1826)	0.18	0.39	-	+
	Sepia officinalis (Linnaeus 1758)	0.08	0.35	-	+
	Loligo vulgaris (Lamarck 1798)	0.13	0.52	-	+
Gastropoda	*Berthellina citrina* (Rüppel & Leuckart 1828)	0.13	0.65	-	+
	Bolma rugosa (Linnaeus 1767)	0.15	0.58	-	+
	Brachystomia sp. (Monterosato 1884)	0.33	0.76	+	+
	Cerithium spp. (Bruguière 1789)	18.1	12.72	+	+
	Charonia tritonis (Linnaeus 1758)	0.13	0.33	+	+
	Conus ventricosus (Gmelin 1791)	4.53	3.99	+	+
	Elysia timida (Risso 1818)	6.85	9.11	+	+
	Fasciolaria lignaria (Linnaeus 1758)	1.85	1.78	+	+
	Flabellina affinis (Gmelin 1791)	0.03	0.16	+	-
	Gibbula racketti (Payraudeau 1826)	0.6	1.53	+	+
	Haliotis tuberculata lamellosa (Lamarck 1822)	2.18	2.31	+	+
	Hexaplex trunculus (Linnaeus 1758)	7	6.94	+	+
	Naticidae sp. (Guilding 1834)	0.08	0.35	-	+
	Patella spp. (Linnaeus 1758)	27.03	24.23	+	+
	Phorcus turbinatus (Born 1778)	3.75	3.64	+	+
	Pyrene scripta (Linnaeus 1758)	87.25	84.01	+	+
	Thuridilla hopei (Vérany 1853)	0.03	0.16	+	-
	Tonna galea (Linnaeus 1758)	0.28	0.68	+	+
Polyplacophora	*Rhyssoplax olivacea* (Spengler 1797)	2.25	4.12	+	+
Total **5**	**60**			**44**	**57**

II Species photographs

Figure 1: Photographs of macroinvertebrate species discussed in the report, presented in alphabetical order. (A) *Actinia equina,* (B) *Astropecten spinulosus*, (C) *Balanophyllia europaea,* (D) *Bonellia viridis*, (E) *Calcinus tubularis*, (F) *Callistoctopus macropus,* (G) *Cerithium sp.,* (H) *Chama pacifica,* (I) *Charonia tritonis,* (J) *Clibinarius erythropus,* (K) *Conus ventricosus,* (L) *Dardanus arrosor.* Photo credits: Janika Sander.

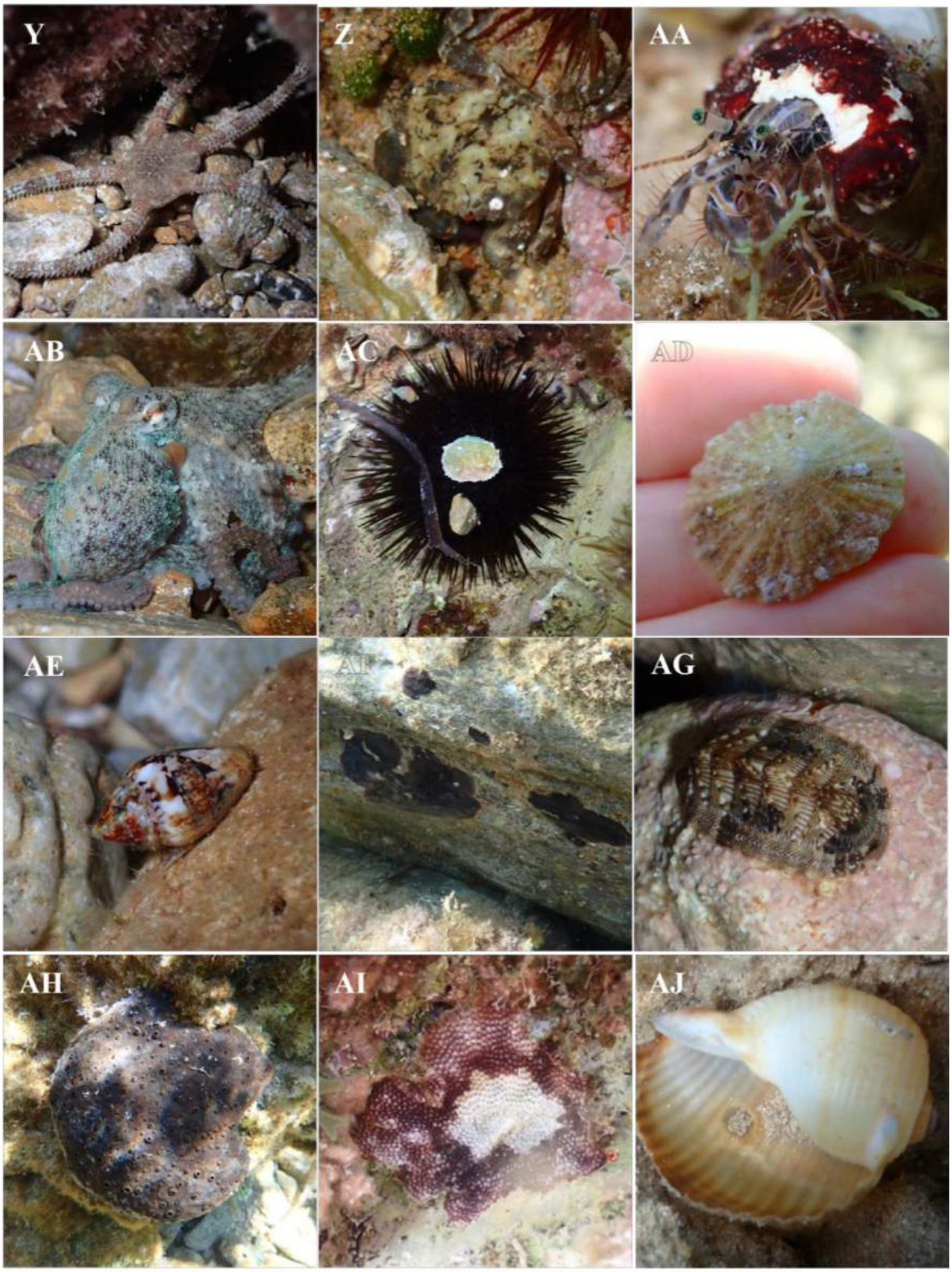

Figure 1: continued. (Y) *Ophioderma longicauda,* (Z) *Pachygraspus marmoratus,* (AA) *Pagurus anachoetus,* (AB) *Octopus vulgaris,* (AC) *Paracentrotus lividus,* (AD) *Patella sp.,* (AE) *Pyrene scripta,* (AF) *Reptadeonella violacea*, (AG) *Rhyssoplax olivacea,* (AH) *Sarcotragus spinulosus,* (AI) *Schizobrachiella sanguinea*, (AJ) *Tonna galea.* Photo credits: Janika Sander.

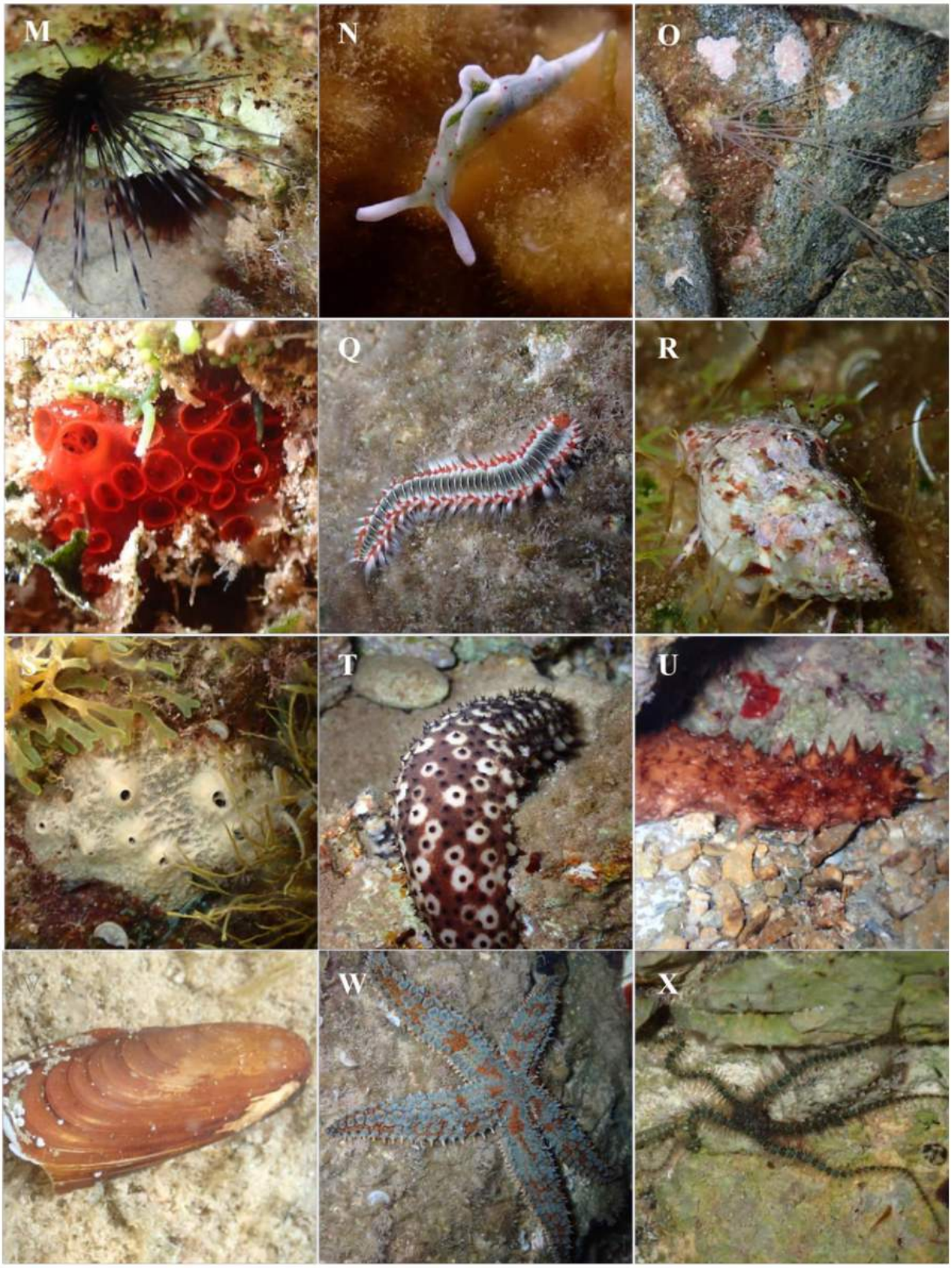

Figure 1: continued. (M) *Diadema setosum,* (N) *Elysia timida,* (O) *Eupolymnia nebulosa,* (P) *Hamigera hamigera,* (Q) *Hermodice caranculata,* (R) *Hexaplex trunculus,* (S) *Ircinia variabilis,* (T) *Holothuria forskali,* (U) *Holothuria tubulosa*, (V) *Lithophaga lithophaga,* (W) *Marthasterias glacialis*, (X) *Ophiothrix fragilis.* Photo credits: Janika Sander.

Bibliography

Aguzzi J, Company JB, Costa C, Matabos M, Azzurro E, Mánuel A, Menesatti P, Sardá F, Canals M, Delory E, et al. 2012. Challenges to the assessment of benthic populations and biodiversity as a result of rhythmic behaviour: Video solutions from cabled observatories. *Oceanogr Mar Biol An Annu Rev. 50*(January):235–286. doi:10.1201/b12157-7.

Alexander BE, Liebrand K, Osinga R, Van Der Geest HG, Admiraal W, Cleutjens JPM, Schutte B, Verheyen F, Ribes M, Van Loon E, et al. 2014. Cell turnover and detritus production in marine sponges from tropical and temperate benthic ecosystems. Webster N, editor. *PLoS One. 9*(10):e109486. doi:10.1371/journal.pone.0109486.

Amon RMW, Herndl GJ. 1991. Deposit Feeding and Sediment: I. Interrelationship Between Holothuria tubulosa (Holothurioida, Ech'modermata) and the Sediment Microbial Community. *Mar Ecol. 12*(2):163–174. doi:10.1111/j.1439-0485.1991.tb00250.x.

Antoniadou C, Vafidis D. 2009. Updated distribution of the holothuroid Synaptula reciprocans (Forskal, 1775) in the Mediterranean: Does it follow shallow-water circulation patterns? *Aquat Invasions. 4*(2):361–363. doi:10.3391/ai.2009.4.2.9.

Apostolopoulou E, Pantis JD. 2009. Conceptual gaps in the national strategy for the implementation of the European Natura 2000 conservation policy in Greece. *Biol Conserv. 142*(1):221–237. doi:10.1016/j.biocon.2008.10.021.

Archipelagos Institute. 2021. Coastal macroinvertebrates in the Aegean Sea [unpublished manuscript].

Aréchiga H, Fernéndez-Quiróz F, de Miguel FF, Rodríguez-Sosa L. 1993. The circadian system of crustaceans. *Chronobiol Int. 10*(1):1–19. doi:10.3109/07420529309064477.

Barua M. 2011. Mobilizing metaphors: the popular use of keystone, flagship and umbrella species concepts. *Biodivers Conserv. 20*:1427–1440. doi: 10.1007/s10531-011-0035-y.

Beaumont NJ, Jones L, Garbutt A, Hansom JD, Toberman M. 2014. The value of carbon sequestration and storage in coastal habitats. *Estuar Coast Shelf Sci. 137*(1):32–40. doi:10.1016/j.ecss.2013.11.022.

Bell JJ. 2008. The functional roles of marine sponges. *Estuar Coast Shelf Sci. 79*(3):341–353. doi:10.1016/j.ecss.2008.05.002.

Benke AC, Huryn AD. 2006. Secondary production of macroinvertebrates. In Methods in Stream Ecology. 2nd ed. Eds. Hauer FR, Lamberti GA. Burlington: Academic Press. 691-710.

Bernard ATF, Götz A, Kerwath SE, Wilke CG. 2013. Observer bias and detection probability in underwater visual census of fish assemblages measured with independent double-observers. *J Exp Mar Bio Ecol. 443*:75–84. doi:10.1016/j.jembe.2013.02.039.

Béthous JP. 1979. Budgets of the Mediterranean sea. Their dependance on the local climate and on the characteristics of the Atlantic waters. *Oceanol Acta. 2*:157–163.

Blackmon DC, Valentine JF. 2013. Recurring nocturnal benthic emergence along the coral reef-seagrass interface in the Florida Keys National Marine Sanctuary: Evidence of a possible novel prey escape response. *J Exp Mar Bio Ecol. 448*:220–227. doi:10.1016/j.jembe.2013.07.012.

Boletzky S V., Fuentès M, Offner N. 2001. First record of spawning and embryonic development in Octopus macropus (Mollusca: Cephalopoda). *J Mar Biol Assoc United Kingdom. 81*(4):703–704. doi:10.1017/S002531540100443X.

Brewin PE, Brown J, Brickle P. 2016. Diurnal variation of fish and macrobenthic invertebrate community structure in an isolated oceanic island of the South Atlantic. *J Mar Biol Assoc United Kingdom. 96*(3):737–747. doi:10.1017/S0025315415000892.

Brock VE. 1954. A Preliminary Report on a Method of Estimating Reef Fish Populations. *J Wildl Manage. 18*(3):297. doi:10.2307/3797016.

Bronstein O, Georgopoulou E, Kroh A. 2017. On the distribution of the invasive long-spined echinoid Diadema setosum and its expansion in the Mediterranean Sea. *Mar Ecol Prog Ser. 583*:163–178. doi:10.3354/meps12348.

Brown ER, Piscopo S, De Stefano R, Giuditta A. 2006. Brain and behavioural evidence for rest-activity cycles in Octopus vulgaris. *Behav Brain Res. 172*(2):355–359. doi:10.1016/j.bbr.2006.05.009.

Brusca RC, Moore W, Shuster SM. 2016. Invertebrates. Third Edition. *Invertebrates.*:193–196, 347.

Bumbeer J, Cattani AP, Chierigatti NB, Da Rocha RM. 2016. Biodiversity of benthic macroinvertebrates on hard substrates in the currais marine protected area, in southern Brazil. *Biota Neotrop. 16*(4). doi:10.1590/1676-0611-BN-2016-0246.

Cardoso P, Borges PAV, Triantis KA, Ferrández MA, Martín JL. 2011. Adapting the IUCN Red List criteria for invertebrates. *Biol Conserv. 144*(10):2432–2440. doi:10.1016/j.biocon.2011.06.020.

Chainho P, Costa JL, Chaves ML, Dauer DM, Costa MJ. 2007. Influence of seasonal variability in benthic invertebrate community structure on the use of biotic indices to assess the ecological status of a Portuguese estuary. *Mar Pollut Bull. 54*(10):1586–1597. doi:10.1016/j.marpolbul.2007.06.009.

Cheal AJ, Thompson AA. 1997. Comparing visual counts

of coral reef fish: Implications of transect width and species selection. *Mar Ecol Prog Ser. 158*(1):241–248. doi:10.3354/meps158241.

Chintiroglou C, Antoniadou C, Vafidis D, Koutsoubas D. 2005. A review on the biodiversity of hard substrate invertebrate communities in the Aegean Sea. *Mediterr Mar Sci. 6*(2):51–62. doi:10.12681/mms.185.

Coll M, Piroddi C, Steenbeek J, Kaschner K, Lasram FBR, Aguzzi J, Ballesteros E, Bianchi CN, Corbera J, Dailianis T, et al. 2010b. The biodiversity of the Mediterranean Sea: Estimates, patterns, and threats. *PLoS One. 5*(8). doi:10.1371/journal.pone.0011842.

Collen B, Bohm M, Kemp R, Baillie JEMJEM, Bohn M, Kemp R, Baillie JEMJEM, Böhm M, Kemp R, Baillie JEMJEM. 2012. Spineless status and trends of the world's invertebrates.

Collier KJ, Probert PK, Jeffries M. 2016. Conservation of aquatic invertebrates: concerns, challenges and conundrums. *Aquat Conserv Mar Freshw Ecosyst. 26*(5):817–837. doi:10.1002/aqc.2710.

Costa Leal M, Puga J, Serôdio J, Gomes NCM, Calado R. 2012. Trends in the discovery of new marine natural products from invertebrates over the last two decades - where and what are we bioprospecting? *PLoS One. 7*(1). doi:10.1371/journal.pone.0030580.

Costanza R, D'Arge R, de Groot R, Farber S, Grasso M, Hannon B, Limburg K, Naeem S, O'Neill R V., Paruelo J, et al. 1997. The value of the world's ecosystem services and natural capital. LK - https://www-nature-com.ezproxy.royalroads.ca/articles/387253a0.pdf.

[CoE] Council of Europe. 1979. Convention on the Conservation of European Wildlife and Natural Habitats Convention relative à la conservation de la vie sauvage et du milieu naturel de l ' Europe.

[CoE] Council of Europe. 2006. Group of Specialists for the European Strategy on Invertebrates. Strasbourg. https://rm.coe.int/168074617c.

[CoE] Council of Europe. 2018. Annexe II - Espèces de faune strictement protégées. Conv Berne Relat àla Conserv la vie Sauvag du milieu Nat l'Europe.(104).

Dee LE, Witman JD, Brandt M. 2012. Refugia and top-down control of the pencil urchin Eucidaris galapagensis in the Galápagos Marine Reserve. *J Exp Mar Bio Ecol. 416–417*(April):135–143. doi:10.1016/j.jembe.2012.02.016.

De Goeij JM, Van Oevelen D, Vermeij MJA, Osinga R, Middelburg JJ, De Goeij AFPM, Admiraal W. 2013. Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science (80-). 342*(6154):108–110. doi:10.1126/science.1241981.

Dunham A, Archer SK, Davies SC, Burke LA, Mossman J, Pegg JR, Archer E. 2018. Assessing condition and ecological role of deep-water biogenic habitats: Glass sponge reefs in the Salish Sea. *Mar Environ Res. 141*:88–99. doi:10.1016/j.marenvres.2018.08.002.

Dupont C, Belin A, Moreira G, Cochrane S, Wilson L, Emblow C, Kater B, Clercs S Des, Parr W, Visage C Le, et al. 2014. Article 12 Technical Assessment of the MSFD 2012 obligations: The United Kingdom.

Edgar GJ, Moverley J, Barrett NS, Peters D, Reed C. 1997. The conservation-related benefits of a systematic marine biological sampling programme: The Tasmanian reef bioregionalisation as a case study. *Biol Conserv. 79*(2–3):227–240. doi:10.1016/S0006-3207(96)00095-X.

Eisenhauer N, Bonn A, A. Guerra C. 2019. Recognizing the quiet extinction of invertebrates. *Nat Commun. 10*(1):1–3. doi:10.1038/s41467-018-07916-1.

[EC] European Commission. 2010. COMMISSION DECISION of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters. *Comm Decis.*(2008):11.

[EC] European Commission. 2018. Report from the Commission to the European Parliament and the Council: Assessing Member States´programme of measures under the Marine Strategy Framewok Directive. Brussels.

[EEA] European Environment Agency. 2020. Natura 2000 - Standard Data Form. [accessed 2021 Jun 4]. https://natura2000.eea.europa.eu/Natura2000/SDF.aspx?site=GR4210010.

[EU] European Union. 2008. Directive 2008/56/ECDIRECTIVE 2008/56/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). [accessed 2021 Jun 4]. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0056&from=EN.

[EU] European Union. 2013. OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 December 2013 on the establishment of a Programme for the Environment and Climate Action (LIFE) and repealing Regulation (EC) No 614/2007 (Text with EEA relevance).

Fingerman M. 1957. Lunar Rhythmicity in Marine Organisms Author (s): Milton Fingerman Source: The American Naturalist , May - Jun ., 1957 , Vol . 91 , No . 858 (May - Jun ., 1957), Published by: *The University of Chicago Press for The American Society of Naturalists S. 91*(858):167–178.

Francour P. 1997. Predation on Holothurians: A Literature Review. *Invertebr Biol. 116*(1):52. doi:10.2307/3226924.

Gabrié C. Bissery C., Crochelet E., Meola B., Webster C., Claudet J., Chassanite A., Marinesque S., LE, Robert P. Quod C. GM, Gabrié C., Lagabrielle E., Bissery C., Crochelet E., Meola B., Webster C., Claudet J., Chassanite A., Marinesque S., Robert P. GM. 2012. The Status of Marine Protected Areas in the Mediterranean Sea.

Giménez-Casalduero F, Muniain C, González-Wangüemert M, Garrote-Moreno A. 2011. Elysia timida (Risso, 1818) three decades of research. *Anim Biodivers Conserv. 34*(1):217–227.

Giorgi F. 2006. Climate change hot-spots. *Geophys Res Lett. 33*(8):L08707. doi:10.1029/2006GL025734.

Gold DA, Grabenstatter J, De Mendoza A, Riesgo A, Ruiz-Trillo I, Summons RE. 2016. Sterol and genomic analyses validate the sponge biomarker hypothesis. *Proc Natl Acad Sci U S A. 113*(10):2684–2689. doi:10.1073/pnas.1512614113.

[MEGR] Greek Ministry of the Environment Energy & Climate Change. 2014. National Biodiversity Strategy & Action Plan. *Athens.* [accessed 2021 Feb 19]. http://www.ypeka.gr.

Habitats Directive. 1992. COUNCIL DIRECTIVE 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora.

Hammond LS. 1982. Deposit-feeding Holothurians and Echinoids (Echinodermata). From a shallow back-reef lagoon. *Discovery. 32*(2):549–571.

Hardin PE, Hall JC, Rosbash M. 1990. Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. *Nature. 343*(6258):536–540. doi:10.1038/343536a0.

Hassan W, Edrada R, Ebel R, Wray V, Proksch P. 2004. New Alkaloids from the Mediterranean Sponge Hamigera hamigera. *Mar Drugs. 2*(3):88–100. doi:10.3390/md203088.

Heywood VH. 1995. Global biodiversity assessment. *Cambridge University Press, Cambridge.*

Hinojosa IA, Zapata-Hernández G, Fowles AE, Gaymer CF, Stuart-Smith RD. 2020. The awakening of invertebrates: The daily dynamics of fishes and mobile invertebrates at Rapa Nui's multiple use marine protected area. *Aquat Conserv Mar Freshw Ecosyst.* 1–14. doi:10.1002/aqc.3423.

Holt RD. 2009. Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proc Natl Acad Sci U S A. 106*(SUPPL. 2):19659–19665. doi:10.1073/pnas.0905137106.

Hutcheson K. 1970. A test for comparing diversities based on the shannon formula. *J Theor Biol. 29*(1):151–154. doi:10.1016/0022-5193(70)90124-4.

Hutchinson GE. 1957. Concluding Remarks. *Cold Spring Harb Symp Quant Biol. 22*(0):415–427. doi:10.1101/sqb.1957.022.01.039.

Ignatiades L. 1998. The productive and optical status of the oligotrophic waters of the Southern Aegean Sea (Cretan Sea), Eastern Mediterranean. [accessed 2021 May 25]. https://academic.oup.com/plankt/article/20/5/985/-1462310.

IPCC. 2019. The Ocean and Cryosphere in a Changing Climate. A Special Report of the Intergovernmental Panel on Climate Change. [accessed 2021 Jun 3]. https://www.ipcc.ch/srocc/chapter/summary-for-policymakers/.

Itoh MT, Shinozawa T, Sumi Y. 1999. Circadian rhythms of melatonin-synthesizing enzyme activities and melatonin levels in planarians. *Brain Res. 830*(1):165–173. doi:10.1016/S0006-8993(99)01418-3.

IUCN. 2021. Table 1a: Number of species evaluated in relation to the overall number of described species, and number of threatened species by major groups of organisms. *Red List Summ Stat.*(July):1. https://www.iucnredlist.org/resources/summary-statistics#Figure 2.

Jones CG, Lawton JH, Shachak M. 1994. Organisms as Ecosystem Engineers. *Oikos. 69*(3):373. doi:10.2307/3545850.

Katsanevakis S, Poursanidis D, Issaris Y, Panou A, Petza D, Vassilopoulou V, Chaldaiou I, Sini M. 2011. "Protected" marine shelled molluscs: Thriving in Greek seafood restaurants. *Mediterr Mar Sci. 12*(2):429–438. doi:10.12681/mms.42.

Koutsoubas D, Galinou-Mitsoudi S, Katsanevakis S, Leontarakis P, Metaxatos A, Zenetos A. 2007. Bivalve and gastropo d mollusks of commercial interest for human consumption in the Hellenic Seas. *State Hell Fish.*(May 2014):70–84.

Kronfeld-Schor N, Dayan T. 2003. Partitioning of Time as an Ecological Resource. *Annu Rev Ecol Evol Syst. 34*:153–181. doi:10.1146/annurev.ecolsys.34.011802.132435.

Kuba M, Meisel D, Byrne R, Griebel U, Mather J. 2003. Looking at play in Octopus vulgaris. Warnke K, Keupp H, Boletzky S, editors. *Berliner Palaeontol Abhandlungen. 3*:163–169.

Lincoln Smith M. 1988. Effects of observer swimming speed on sample counts of temperate rocky reef fish assemblages. *Mar Ecol Prog Ser.* 43:223–231. doi:10.3354/meps043223.

Lionello P, Malanotte-Rizzoli P, Boscolo R, Alpert P, Artale V, Li L, Luterbacher J, May W, Trigo R, Tsimplis M, et al. 2006. The Mediterranean climate: An overview of the main characteristics and issues. *Dev Earth Environ Sci. 4*(C):1–26. doi:10.1016/S1571-9197(06)80003-0.

MacFarlane GR, King SA. 2002. Observer presence influences behaviour of the semaphore crab, Heloecious cordiformis. *Anim Behav. 63*(6):1191–1194. doi:10.1006/anbe.2002.3016.

Mammola S, Riccardi N, Prié V, Correia R, Cardoso P, Lopes-Lima M, Sousa R. 2020. Towards a taxonomically unbiased European Union biodiversity strategy for 2030. *Proc R Soc B Biol Sci. 287*(1940):20202166. doi:10.1098/rspb.2020.2166.

Meisel D V., Byrne RA, Kuba M, Mather J, Ploberger W, Reschenhofer E. 2006. Contrasting activity patterns of two related Octopus species, Octopus macropus and Octopus vulgaris. *J Comp Psychol. 120*(3):191–197. doi:10.1037/0735-7036.120.3.191.

[MA] Millennium Ecosystem Assessment. 2005. Ecosystems and Human Well-being: Biodiversity Synthesis. Washington DC. [accessed 2021 Jun 4]. https://www.millenniumassessment.org/documents/-document.354.aspx.pdf.

Neuhäuser M. 2011. Wilcoxon–Mann–Whitney Test. In: International Encyclopedia of Statistical Science. *Springer Berlin Heidelberg*. p. 1656–1658.

OECD. 2021. Fisheries and Aquaculture in Greece. In: OECD Review of Fisheries Country Notes. [accessed 2021 May 23]. https://www.oecd.org/agriculture/topics/fisheries-and-aquaculture/documents/report_cn_fish_grc.pdf.

Officer C, Smayda T, Mann R. 1982. Benthic Filter Feeding: A Natural Eutrophication Control. *Mar Ecol Prog Ser. 9*:203–210. doi:10.3354/meps009203.

Ory NC, Dudgeon D, Duprey N, Thiel M. 2014. Effects of predation on diel activity and habitat use of the coral-reef shrimp Cinetorhynchus hendersoni (Rhynchocinetidae). *Coral Reefs. 33*(3):639–650. doi:10.1007/s00338-014-1163-0.

Ostroumov SA. 2005. Suspension-Feeders as Factors Influencing Water Quality in Aquatic Ecosystems. In: The Comparative Roles of Suspension-Feeders in Ecosystems. *Springer-Verlag*. p. 147–164. doi: 10.1007/1-4020-3030-4_9.

Pagès Escolà M. 2019. New insights into the ecology and conservation of bryozoans: from global diversity patterns to the responses to anthropogenic stressors in the Mediterranean Sea. Universitat de Barcelona. [accessed 2021 Jun 6]. https://www.openchannels.org/sites/default/files/literature/new_insights_into_the_ecology_and_conservation_of_bryozoans.pdf.

Palialexis A, Cardoso AC, Somma F. 2018. JRC's reference lists of MSFD species and habitats.

Pelletreau KN, Bhattacharya D, Price DC, Worful JM, Moustafa A, Rumpho ME. 2011. Sea slug kleptoplasty and plastid maintenance in a metazoan. *Plant Physiol. 155*(4):1561–1565. doi:10.1104/pp.111.174078.

Peres R, Amaral FG, Marques AC, Neto JC. 2014. Melatonin production in the sea star Echinaster Brasiliensis (Echinodermata). *Biol Bull. 226*(2):146–151. doi:10.1086/BBLv226n2p146.

Poulos SE, Drakopoulos PG, Collins MB. 1997. Seasonal variability in sea surface oceanographic conditions in the Aegean Sea (Eastern Mediterranean): An overview. *J Mar Syst. 13*(1–4):225–244. doi:10.1016/S0924-7963(96)00113-3.

Prather CM, Pelini SL, Laws A, Rivest E, Woltz M, Bloch CP, Del Toro I, Ho CK, Kominoski J, Newbold TAS, et al. 2013. Invertebrates, ecosystem services and climate change. *Biol Rev. 88*(2):327–348. doi:10.1111/brv.12002.

R Core Team. 2020. A language and environment for statistical computing. R Foundation for Statistical Computing. [accessed 2021 Jun 6]. https://www.r-project.org/index.html.

Reid T. 2000. The Other 99%. The Conservation and Biodiversity of Invertebrates. *Pacific Conserv Biol. 6*(2):178. doi:10.1071/pc000178.

Rife GS. 2014. Impacts of Tsunami events on ecosystem services provided by benthic macro-invertebrate assemblages of marine coastal zones. *Adv Nat Technol Hazards Res. 35*:147–159. doi:10.1007/978-94-007-7269-4_7.

Roenneberg T, Daan S, Merrow M. 2003. The art of entrainment. *J Biol Rhythms. 18*(3):183–194. doi:10.1177/0748730403018003001.

Rosbash M. 2017. Circadian Rhythms and the Transcriptional Feedback Loop. Waltham MA, USA.

Rosbash M, Hall JC. 1989. The molecular biology of circadian rhythms. *Neuron. 3*(4):387–398. doi:10.1016/0896-6273(89)90199-2.

Royston JP. 1982. An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. *Appl Stat. 31*(2):115. doi:10.2307/2347973.

Ryu SH, Jang KH, Choi EH, Kim SK, Song SJ, Cho HJ, Ryu JS, Kim YM, Sagong J, Lee JH, et al. 2012. Biodiversity of marine invertebrates on rocky shores of Dokdo, Korea. *Zool Stud. 51*(5):710–726.

Santinelli C. 2015. DOC in the Mediterranean Sea. In: Biogeochemistry of Marine Dissolved Organic Matter: Second Edition. *Elsevier Inc.* p. 579–608.

Sato T. 1996. Type i and type ii error in multiple comparisons. *J Psychol Interdiscip Appl. 130*(3):293–302. doi:10.1080/00223980.1996.9915010.

Schoener TW. 1974. Resource Partitioning in Ecological Communities Published by: American Association for the Advancement of Science Stable.

Schroeder K, Garcìa-Lafuente J, Josey SA, Artale V, Nardelli BB, Carrillo A, Gačić M, Gasparini G Pietro, Herrmann M, Lionello P, et al. 2012. Circulation of the mediterranean sea and its variability. In: The Climate of the Mediterranean Region. *Elsevier Inc.* p. 187–256.

Seymour AML. 2003. Imaging cardiac metabolism in heart failure: The potential of NMR spectroscopy in the era of metabolism revisited. *Hear Lung Circ. 12*(1):25–30. doi:10.1046/j.1444-2892.2003.00176.x.

Shannon CE. 1948. A Mathematical Theory of Communication. *Bell Syst Tech J. 27*(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.

Shick J, Dunlap W, Chalker B, Banaszak A, Rosenzweig T. 1992. Survey of ultraviolet radiation-absorbing mycosporine-like amino acids in organs of coral reef holothuroids. *Mar Ecol Prog Ser. 90*(December 1992):139–148. doi:10.3354/meps090139.

Sini M, Katsanevakis S, Koukourouvli N, Gerovasileiou V, Dailianis T, Buhl-Mortensen L, Damalas D, Dendrinos P, Dimas X, Frantzis A, et al. 2017. Assembling ecological pieces to reconstruct the conservation puzzle of the aegean sea. *Front Mar Sci. 4*(NOV):347. doi:10.3389/fmars.2017.00347.

Stabili L, Schirosi R, Parisi MG, Piraino S, Cammarata M. 2015. The mucus of Actinia equina (Anthozoa, Cnidaria): An unexplored resource for potential applicative purposes. *Mar Drugs. 13*(8):5276–5296. doi:10.3390/md13085276.

Sullivan KM, Chiappone M. 1992. A comparison of belt quadrat and species presence/absence sampling of stony coral (Scleractinia and Milleporina) and sponges for evaluating species patterning on patch reefs of the central Bahamas. *Bull Mar Sci. 50*(3):464–488.

Tessmar-Raible K, Raible F, Arboleda E. 2011. Another place, another timer: Marine species and the rhythms of life. *BioEssays. 33*(3):165–172. doi:10.1002/bies.201000096.

Thukral AK. 2017. A review on measurement of Alpha diversity in biology. *Agric Res J. 54*(1):1. doi:10.5958/2395-146x.2017.00001.1.

Trani R, Corriero G, de Pinto MC, Mercurio M, Pazzani C, Pierri C, Scrascia M, Longo C. 2021. Filtering activity and nutrient release by the keratose sponge sarcotragus spinosulus schmidt, 1862 (Porifera, demospongiae) at the laboratory scale. *J Mar Sci Eng. 9*(2):1–16. doi:10.3390/jmse9020178.

Tsikliras AC, Dinouli A, Tsalkou E. 2013. Exploitation trends of the Mediterranean and Black Sea fisheries. *Acta Adriat. 54*(2):273–281.

Usseglio P. 2015. Quantifying reef fishes: Bias in observational approaches.

Van Nguyen T, Cho W-S, Kim H, Jung IH, Kim Y, Chon T-S. Inferring community properties of benthic macroinvertebrates in streams using Shannon index and exergy. doi:10.1007/s11707-013-0420-9.

Voultsiadou E. 2005. Sponge diversity in the aegean sea: Check list and new information. *Ital J Zool. 72*(1):53–64. doi:10.1080/11250000509356653.

Voultsiadou E, Koutsoubas D, Achparaki M. 2010. Bivalve mollusc exploitation in Mediterranean coastal communities: An historical approach. *J Biol Res.* 13:35–45.

Voultsiadou E, Vafidis D. 2007. Marine invertebrate diversity in Aristotle's zoology. *Contrib to Zool. 76*(2):103–120. doi:10.1163/18759866-07602004.

Wang X, Wang Q, Yang Y, Yu W. 2020. Comparison of invertebrate diversity in lake waters and their resting eggs in sediments, as revealed by high-throughput sequencing (HTS). *Knowl Manag Aquat Ecosyst.* 2020-January(421):19. doi:10.1051/kmae/2020011.

Webb P. 2019. Introduction to Oceanography. *Online OER textbook.* https://webboceanography.pressbooks.com.

Wheeling RJ, Verde EA, Nestler JR. 2007. Diel cycles of activity, metabolism, and ammonium concentration in tropical holothurians. *Mar Biol. 152*(2):297–305. doi:10.1007/s00227-007-0683-3.

Whittaker RH. 1972. Evolution and Measurement of Species Diversity.

Williams ID, Walsh WJ, Tissot BN, Hallacher LE. 2006. Impact of observers' experience level on counts of fishes in underwater visual surveys. *Mar Ecol Prog Ser. 310*(Kulbicki 1998):185–191. doi:10.3354/meps310185.

Wilsey BJ, Potvin C. 2000. Biodiversity and ecosystem functioning: Importance of species evenness in an old field. *Ecology. 81*(4):887–892. doi:10.1890/0012-9658.

Wilson EO. 1987. The Little Things That Run the World (The Importance and Conservation of Invertebrates).

WoRMS Editorial Board (2021). World Register of Marine Species. Available from https://www.marinespecies.org at VLIZ. [accessed 2021 May 23]. doi:10.14284/170.

Wulff JL. 2006. Ecological interactions of marine sponges. *Can J Zool. 84*(2):146–166. doi:10.1139/z06-019.

Zenetos A, Gofas S, Verlaque M, Çinar ME, García Raso JG, Bianchi CN, Morri C, Azzurro E, Bilecenoglu M, Froglia C, et al. 2010. Alien species in the Mediterranean Sea by 2010. A contribution to the application of European Union's Marine Strategy Framework Directive (MSFD). Part

I. Spatial distribution. *Mediterr Mar Sci*. *11*(2):381–493. doi:10.12681/mms.87.

Zenetos A, Pancucci-Papadopoulou MA, Zogaris S, Papastergiadou E, Vardakas L, Aligizaki K, Economou AN. 2009. Aquatic alien species in Greece (2009): Tracking sources, patterns and effects on the ecosystem. *J Biol Res*. *12*:135–172.

Zenetos A, Vardala-Theodorou E, Alexandrakis C. 2005. Update of the marine Bivalvia Mollusca checklist in Greek waters. *J Mar Biol Assoc United Kingdom*. *85*(4):993–998. doi:10.1017/S0025315405012014.

Sciences

NGC 1052–DF2: The Galaxy Re-sparking the Modified Gravity Debate

A Systematic Literature Review of the Ultra-Diffuse Galaxy, NGC 1052–DF2, and its Significance for the Cold Dark Matter (CDM) and Modified Newtonian Dynamics (MOND) Paradigms

Karla M. Rojas

Supervisor
Dr. Gianfranco Bertone (UvA)
Reader
Dr. Margriet van der Heijden (AUC)

Photographer: Iacob Postavaru

Abstract

This thesis discusses the novel finding that the ultra-diffuse galaxy (UDG) NGC1052–DF2 lacks dark matter in the context of Cold Dark Matter (CDM) and Modified Newtonian dynamics (MOND). At the time of writing, no paper in the literature provided a comprehensive review of this UDG despite its significance for both models and its unusual characteristics. However, understanding the true nature of dark matter and dark-matter-deficient galaxies is vital to determine the formation and evolution of galaxies and, ultimately, the universe's fate. To that aim, this thesis introduces the reader to three brief sections on CDM, MOND, and UDGs, followed by an analysis of NGC1052–DF2 within the modified gravity debate using a systematic literature review. This paper finds that both CDM and MOND are in agreement with observations. However, recent findings concerning the separation of NGC1052–DF2 from another UDG might challenge MOND if the suspicion that one of them is isolated is proven.

Keywords: *NGC 1052–DF2, CDM, MOND, ultra-diffuse galaxy, dark-matter-deficient galaxy*

Suggested citation:
Rojas, K. (2021). NGC 1052–DF2: The galaxy re-sparking the modified gravity debate: A systematic literature review of the ultra-diffuse galaxy, NGC 1052–DF2, and its significance for the cold dark matter (CDM) and modified Newtonian dynamics (MOND) paradigms. *AUC Undergraduate Journal of Liberal Arts and Sciences, 16*(1), 30-45.

Contents

Abbreviations

Cold dark matter (CDM)
Cosmic background radiation (CMB)
External field effect (EFE)
Globular clusters (GCs)
High surface brightness (HBS) galaxies
International System of Units (SI)
Low surface brightness (LSB) galaxies
Matter power spectra (MPS)
Modified Newtonian dynamics (MOND)
Navarro-Frenk-White (NFW) profile
Pisces Stellar Stream (PSS)
Stellar-to-halo mass ratio (SHM)
Strong equivalence principle (SEP)
Surface brightness fluctuation (SBF) technique
Tensor-vector-scalar gravity (TeVeS)
Tidal dwarf galaxies (TDGs)
Tip of the red giant branch (TRGB) technique
Tracer mass estimator (TME) method
Ultra-diffuse galaxies (UDGs)

1 Introduction

Dark matter is perhaps the most prominent open mystery in physics. Thanks to studies of the mass of galaxy clusters (Smith, 1936; Zwicky, 1933) and of galactic rotation curves (Babcock, 1939; Rubin & Ford, 1970), where orbital velocity is plotted as a function of radial distance, it is well-established that ordinary or baryonic mass (i.e., stars, planets, gas) cannot account for all the inferred dynamical mass, which is the total mass that contributes to gravity. This discrepancy became known as the missing mass problem (for a review see Gunn (1980)). Consequently, there have been proposals that dark matter is the most probable solution to this problem. Dark matter consists of a new type of elementary particle. These particles faintly interact with each other and with the rest of matter through only gravitational force.

However, the failure to detect these particles (Crisosto et al., 2020; Ma et al., 2020; SENSEI Collaboration, 2019; XENON Collaboration, 2020) has led some scientists to believe that the problem lies in our understanding of gravity. This has led to the development of several modified gravity theories. While large-scale data better fits dark matter models (Bahcall, 2015) galaxy-scale data favors modified gravity theories (Lisanti et al. 2019). Thus far, no single model has been able to explain all the observational data.

While the debate continues, an intriguing new piece of observational data came into play. In 2018, Pieter van Dokkum and his team reported the first dwarf galaxy, NGC1052–DF2, lacking dark matter. Such a discovery has profound implications for galaxy formation and dark matter models. However, despite its significance, there is no literature review focused on NGC1052–DF2 yet. Hence, this paper aims to analyse NGC1052–DF2 under the two leading paradigms of dark matter and modified gravity—cold dark matter (CDM) and Modified Newtonian dynamics (MOND)—with the objective to attempt to elucidate which model better explains the observational data in this context.

2 Research Context

2.1 The Cold Dark Matter (CDM) model

In the early 1980s, the CDM cosmological model was put forward independently by three teams (Blumenthal, 1982; Bond et al., 1982; Peebles, 1982). It states that cold dark matter is a hypothetical type of particle that moves slowly compared to the speed of light; hence, it is considered 'cold'. It only interacts with baryonic or ordinary matter through gravitational force, and does not emit, absorb or reflect light; it is 'dark'.

In comparison, the hot dark matter model is based on particles that move at relativistic speeds, but some disregard this model due to problems with cosmic structure forming top-down (Primack & Gross, 2001). The latter process refers to large structures (i.e., superclusters) forming first and smaller ones later on through fragmentation; however, this opposes what is inferred from galaxy distribution observations (Tasitsiomi, 2003). One of the biggest challenges in high-energy physics is to find the particle that makes up dark matter. While there are many possibilities, some popular candidates include axions, inert Higgs doublets, sterile neutrinos, and weakly interacting massive particles (WIMPS) (Feng, 2010).

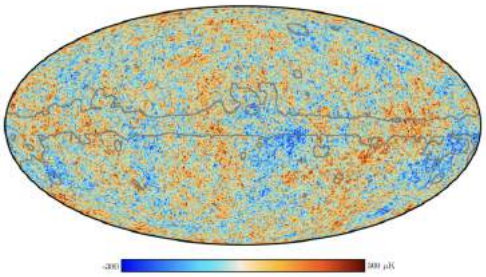

Figure 1. Map showing the temperature anisotropies of the cosmic microwave background (CMB) radiation. The colours represent the temperature measured in microkelvins, and the grey line shows the confidence mask. Source: Planck Collaboration, 2018.

In the CDM framework, as Blumenthal et al. (1984) summarize, it assumed that the universe: 1) originated from the Big Bang, 2) follows the cosmological principle, which holds that space is homogeneous and isotropic on cosmic scales, 3) is almost spatially flat, in which case Euclidean geometry is valid even at large scales, 4) exhibits a cosmological expansion, where the distance between gravitationally unbound parts of the universe is increasing with time, and 5) underwent a period of rapid expansion, known as inflation (Guth, 1981), where it grew by a factor higher than 10^{26} (National Aeronautics and Space Administration, 2010). Consequently, the quantum-mechanical fluctuations present in the early universe were amplified, resulting in the energy density variations observed today on astrophysical scales shown in Figure 1 (Planck Collaboration, 2018). Additionally, Einstein's General Theory of Relativity is deemed as the correct description of gravity.

This model excels at explaining the origin and evolution of large-scale structures in the universe. Galaxies are distributed along filaments and those, in turn, form a percolating network (Shandarin et al., 2010). However, simulations show that baryonic matter processes have almost no impact on large-scale structures (Cui et al., 2017) and are thus unable to produce such networks. Instead, cosmic systems are formed and maintained thanks to the gravitational attraction of CDM particles that start a clustering process. This process begins with small overdensities that prevail over the cosmological expansion and create dark matter halos. Numerical simulations suggest that these halos exhibit a universal density profile often modelled with the NFW (Navarro-Frenk-White) profile, which calculates the density of dark matter

as a function of the radius (Navarro et al., 1997):

$$p(r) = \frac{\delta_c p_c}{\frac{r}{r_s}(1+\frac{r}{r_s})^2}$$

where r_s is the scale radius; δ_c is the characteristic density, which is a dimensionless constant; p_c is the critical density, which refers to the density required to stop cosmological expansion and has a value of $p_c = \frac{3H^2}{8\pi G}$; H is Hubble's constant with a value of 70 $\frac{km}{sMpc}$; and G is the gravitational constant with a value of 6.67 $\frac{m^3}{kgs^2}$.

These initial dark matter halos merge with other halos resulting in the ideal environment for galaxy formation; baryons are trapped in these regions and then forced to cool and collapse within them. Consequently, structures form hierarchically: galaxies formed first, clusters of galaxies followed, and today superclusters are just becoming bound (Pearson & Batuski, 2013). This process assembles larger and larger objects with time and explains the present ongoing galaxy formation. In addition, the CDM model is in impressive agreement with a wide range of physical observations: the primordial abundance of deuterium (Burles et al., 2001), the age of the universe based on the oldest stars (Chaboyer et al., 1998), and the cosmic microwave background (CMB) radiation (Netterfield et al., 2002), which is the remnant electromagnetic radiation from the very early universe. This is particularly relevant, as agreement with physical observation among theories in physics confers validity to the theory as a whole.

Despite its success on large scales, the model has inconsistencies on the galactic and subgalactic scales. In the past, the results of high-resolution cosmological N-body simulations revealed two main issues. Firstly, CDM collapse produces cuspy dark matter halos, whereas observations of galaxy rotation curves hint at a constant density core (Oh et al., 2015). This is known as the cusp-core problem and is still an issue particularly within the dwarf galaxy population (Oman et al., 2015). However, the inclusion of baryonic effects into CDM-only simulations might greatly alleviate the issue (Weinberg et al., 2015 and references therein). Secondly, simulated halos possess a large amount of substructure—hundreds or thousands of subhalos—formed by previous collapses on smaller scales (Moore et al., 1999), yet observations initially found only a few satellite galaxies around larger ones (i.e., around 10 for the Milky Way). This tension was known as the missing satellites problem. However, this is no longer an issue (Kim et al., 2018) thanks to recent astronomical surveys that have counted the correct number of satellites.

2.2 The Modified Newtonian dynamics (MOND) model

In 1983, Milgrom proposed an alternative to the dark matter paradigm motivated by the "large number of ad hoc assumptions concerning the nature of the hidden mass and its distribution in space" (p. 365). While studying galaxies, Milgrom noticed that the difference between the observable baryonic mass and the dynamical mass, which is the total mass predicted using Newton's inverse square law, did not appear in large systems but in systems of low acceleration. Hence, he modified Newton's second law of motion by introducing an interpolating function $\mu(x)$:

$$F = \mu\left(\frac{a}{a_0}\right)a$$

which replaces $F = ma$. In the previous expression, m_g is the body's gravitational mass; a is its gravitational acceleration; a_0 is a new fundamental constant which Milgrom originally determined to have a value of $2x10^{-8}\frac{cm}{s^2}$. Lastly, $\mu(x)$ is the so-called interpolating function, which does not have a determined value but rather must exhibit an asymptotic behaviour $\mu(x)\ x = 1$ when $x \gg 1$ and $\mu(x) = x$ when $x \gg 1$ to be in agreement with Newtonian mechanics and due to astronomical constraints, respectively. The interpolating function can take many forms, but there are two popular formulations (Gentile et al., 2011).

$$\mu\left(\frac{a}{a_0}\right) = \frac{1}{1+\frac{a_0}{a}}$$

is known as the simple interpolating function. In comparison, the standard interpolating function has a root square, and one of the terms in the denominator is quadratic:

$$\mu\left(\frac{a}{a_0}\right) = \sqrt{\frac{1}{1+(\frac{a_0}{a})^2}}$$

In both of these two formulations, the fundamental constant, a_0, delimits the boundary between Newtonian and MOND dynamics (Milgrom, 2014): at high accelerations, Newtonian's dynamics are recoverable (i.e., in the limit of $a_0 \to 0$), and at low accelerations (i.e., $a \ll a_0$), the so-called deep-MOND regime is entered, where dynamics are space and time scale invariant—that is, the system remains unchanged when multiplied by suitable scaling factors. It is worth noting that the original MOND model is only applicable in fields with weak relativistic effects.

A unique property of the nonlinearity of the MOND model is its external field effect (EFE), which goes against general relativity's strong equivalence principle (SEP). The latter holds that while there are no isolated bodies in the Universe, the internal dynamics of a subsystem are independent of any external fields of which the subsystem might be part; that is, it effectively decouples from its parent's field. In contrast, in a MONDian universe, the subsystem does not always go through such decoupling. MONDian effects are visible if the absolute value of the internal gravity g and that of the external gravity g_e is less than a_0. There are two scenarios here (Famaey, 2012). Firstly, if g is bigger than g_e the system exhibits the expected MOND effects. If instead g_e is bigger than g, the system is Newtonian but has a renormalized gravitational constant G_{norm}. Finally, as previously mentioned, if g_e is bigger than both a_0 and g, the system is simply Newtonian.

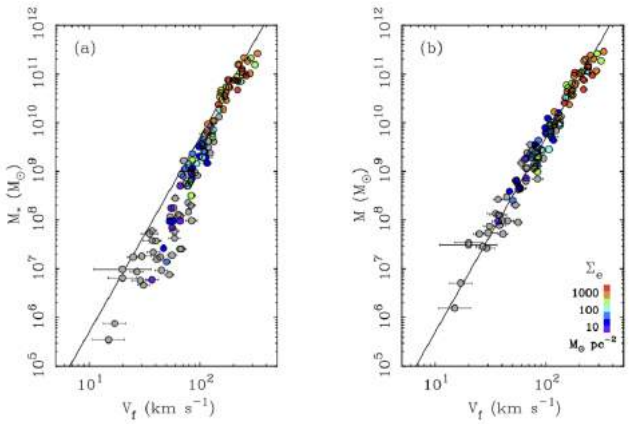

Figure 2. The Tully-Fisher relation. The first panel plots rotational velocity versus stellar mass and the second panel plots rotational velocity against total baryonic mass. Blue represents a low surface brightness value and red a high one, while grey means there is no data to generate such a classification. Source: McGaugh, 2020.

Arguably, MOND is unrivalled at explaining the dynamics of spiral and elliptical galaxies. The model can successfully explain key observations like the galaxies' rotation curves (Sanders & McGaugh, 2002) and the fitted mass-to-light ratio $\frac{M^*}{L}$ (Sanders & Verheijen, 1998), which refers to the ratio between the total mass of a volume and its luminosity. Furthermore, as Figure 2 shows, the model fits well with the empirical correlation between luminosity and rotational velocity of spiral galaxies, known as the Tully-Fisher relation (Tully & Fisher, 1977). Notably, MOND predicted the line shown in Figure 2 before the data even existed (McGaugh, 2020), illustrating that this model is quite strong at the galactic scale.

Nevertheless, the original formulation of MOND did not constitute a complete theory. As Milgrom pointed out, people should regard his modification as an effective working equation. The model violated some of the fundamental laws of physics—the conservation of momentum, angular momentum, and energy—making it incompatible with general relativity (Bertone & Hooper, 2018). However, this changed when Bekenstein (2004) proposed a relativistic generalisation of MOND known as Tensor-vector-scalar gravity (TeVeS). This relativistic theory could explain conservation laws and gravitational lensing, and could reproduce MOND results in the limits of a weak gravitational field.

Despite its relativistic reformulation, MOND still faces crucial hurdles on the galaxy cluster scale. For instance, it is unable to explain the phenomenon seen in the Bullet cluster (1E 0657-56), where the location of baryonic matter does not match the distribution of matter obtained through gravitational lensing (Clowe et al., 2004). If the cluster was composed only of baryonic matter, as MOND proposes, these two locations should be identical. Milgrom (2007) informally argued that MOND could still explain the Bullet cluster if undetected baryonic matter (i.e., faint stars, cold gas clouds, neutrinos) were present in the same proportion as visible matter. However, most of the scientific community saw this as a significant strain on the theory (Paraficz et al., 2016) and considered this to be a less likely explanation to the missing mass problem.

However, the theory recently regained attention thanks to Skordis and Zlosnik's (2020) promising relativistic MOND reformulation. It reproduces TeVeS theory and, in addition, agrees with important observables like the cosmic background radiation (CMB) and the matter power spectra (MPS), which describe the difference between local and average density in the universe. Their theory still needs to be developed further since they deem it a "limiting case of a more fundamental theory" (p. 6). Nonetheless, it shows that the debate between dark matter and modified gravity is still unsettled.

2.3 Ultra-diffuse galaxies (UDGs)

In the last few decades, we began to explore an unfathomed world: one of low surface brightness (LSB) galaxies. These objects are vital to understanding galaxy evolution since they dominate the galaxy number density (Martin et al., 2019). Furthermore, as Kaviraj (2020) highlighted, our current physical models might be biased towards a smaller sample of high surface brightness (HBS) galaxies. Van den Bergh published the first catalogue of LSB objects in 1959. This was limited to low-mass galaxies primarily located in the Local Group—the galaxy group that contains the Milky Way and Andromeda. However, thanks to deep-wide surveys like 2MASS (Skrutskie et al., 2006) or UKIDSS (Lawrence et al., 2007), we are increasingly discovering LSB objects of different mass ranges within cosmological distances.

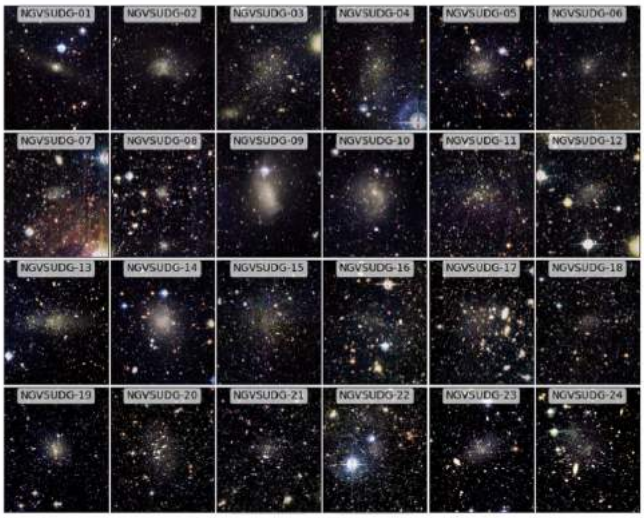

Figure 3. Ultra-diffuse galaxies (UDGs) from the Next Generation Virgo Cluster Survey (NGVS). These images were taken with the Megacam, the wide-field imaging camera of the Canada France Hawaii Telescope. Source: Lim et al., 2020.

Of particular interest for this paper are ultra-diffuse galaxies (UDG), shown in Figure 3. These are a subtype of LSB galaxies that have very low surface brightness ($\mu > 24.5\ \frac{mag}{arcsec^2}$) with masses like the ones of dwarf galaxies ($M = 10^8_\odot$) but with sizes comparable to giants, like the Milky Way ($R_e \sim\ 1.5 - 5kpc$) (van Dokkum et al., 2015). These systems have served to study the mass distribution of

galaxies and their stellar feedback processes (Di Cintio et al., 2017). These galaxies have intrigued scientists due to their complex formation mechanisms as they have been detected in both inside (Mancera Piña et al., 2018) and outside galaxy clusters (Román & Trujillo, 2017). Recently, interest surged with van Dokkum et al.'s (2018a) claim that a UDG lacked dark matter. This claim further puzzled the scientific community because it was unclear how a system dependent on dark matter for its formation observed a depletion of non-baryonic matter.

The formation mechanism of UDGs, however, is still not well-understood. There are at least three possible scenarios outlined in the literature. On the one hand, van Dokkum et al., (2015) put forward the idea that UDGs are L* galaxies (galaxies with a stellar-mass similar to the Milky Way's), which have massive dark matter halos that failed to retain their stellar components. Under this hypothesis, UDGs can survive in very dense environments (e.g., close to cluster centres) because of their high mass. On the other hand, Amorisco & Loeb (2016) proposed that UDGs follow the standard model of disk formation, where the size of galaxies is determined by the spin of their dark matter halo. This process resembles that which LSB dwarfs undergo but with different initial conditions (i.e., a high angular momentum content), which leads to an extended disk. Di Cintio et al.'s (2017) simulations supported this hypothesis. Lastly, Román & Trujillo (2017) suggested that UDGs form outside clusters and become part of such systems through the infall of galaxy groups.

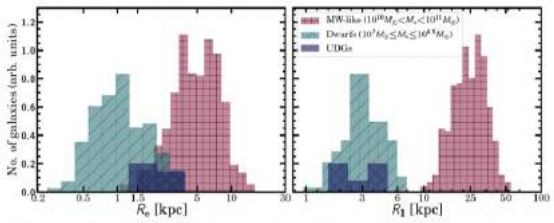

Figure 4. Histograms depicting the size distribution of ultra-diffuse galaxies (UDGs), dwarf galaxies, and Milky Way-like galaxies. The left panel uses the effective radius. In comparison, the right panel uses the radius based on gas density. Consequently, the overlap between UDGs and MW-like galaxies disappears, and UDGs are only linked to dwarfs. Source: Chamba et al., 2019.

Whilst the previously outlined scenarios all operated under the assumption that UDGs are Milky Way sized, Chamba et al. (2019) recently challenged this notion. They found that most UDGs are in the size scale of dwarf galaxies (i.e., 10 times smaller than the Milky Way). These discrepancies stem from measuring galaxies based on the effective radius, R_e, which is the radius that captures half of the total light emitted by a system. The team argued that the effective radius is unable to reach the physical boundaries of a galaxy since the measured radius depends on how light is concentrated in the system. Hence, using the effective radius would lead to unfair comparisons with UDGs considering that a fainter configuration produces a higher radius. The team measured the galaxy size based on the gas density threshold required for star formation to avoid this problem. The threshold reflects the physical boundaries of a galaxy since outside the threshold the gas density is so low that it cannot start stellar formation, and can thus be considered the 'boundary' or 'edge' of the galaxy. As Figure 4 shows, the overlap between UDGs and Milky Way-like galaxies disappears. Under these circumstances, van Dokkum et al.'s (2015) formation mechanism seems less plausible. Nevertheless, as Lim et al. (2020) emphasise, more research and data collection are needed before any option is ruled out as the data sets used to study UDGs so far have been incomplete and heterogeneous.

3 Methodology

This thesis used a systematic literature review and the ultra-diffuse galaxy NGC1052–DF2 as a case study to analyse the predictions of the MOND and CDM paradigms. The research question was: Which model can better describe the observational data and behaviour of NGC1052–DF2, MOND, or CDM? As Figure 5 displays, the journals considered were: Nature Astronomy, The Astrophysical Journal, The Astrophysical Journal Letters, Research notes of the AAs, Astronomy & Astrophysics, and Monthly Notices of the Royal Astronomical Society. Searching for the terms "NGC1052–DF2," "dark matter," "MOND," "CDM," "Standard Cosmology," and "ultra-diffuse galaxy" in the combinations outlined in Figure 5 generated 31 results. After an eligibility screening, only the 28 papers explicitly discussing NGC1052–DF2 were considered.

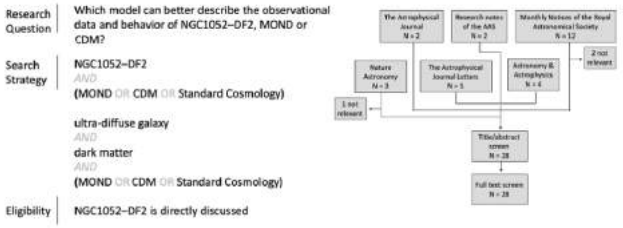

Figure 5. On the left, the research question, search strategy, and eligibility criteria of the systematic literature review. On the right, a flow chart of the eligibility screening. Source: Self-made.

4 Discussion

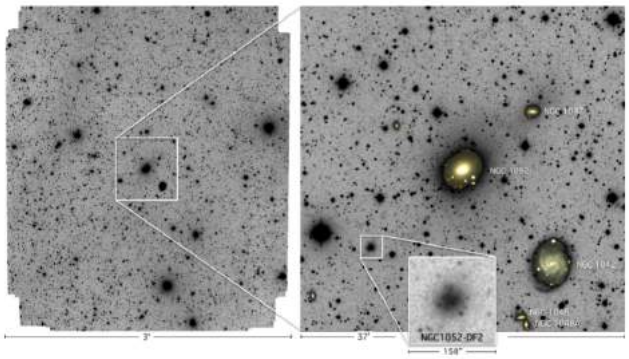

Figure 6. Image of the NGC 1052 group taken with the Dragonfly Telephoto Array. Source: van Dokkum et al., 2018a.

Van Dokkum et al. (2018a) reanimated the debate between the cold dark matter (CDM) and modified Newtonian

dynamics (MOND) paradigms when they announced that the UDG NGC 1052–DF2, shown in Figure 6, lacked dark matter. This section will provide an in-depth analysis of the characteristics of the UDG system, which will serve as a case study to compare and contrast the predictions of both models.

First, however, I present a simplified derivation of van Dokkum et al.'s (2018a) result. The team conducted a kinematic study, which consists of measuring the motions of celestial objects, based on ten globular clusters— spherical gravitationally bound groups of stars that orbit a galaxy—to obtain a velocity dispersion value of less than 10.5 $\frac{km}{s}$. If a galaxy is approximated as a sphere, the velocity dispersion can be used rather than the velocity to derive the so-called dynamical mass of the system. While the team used the sophisticated tracer mass estimator (TME) method, I will provide a simple order of magnitude estimate of the total mass using the virial theorem:

$$2T + U = 0$$

which is applicable only to gravitationally bound systems. Here T denotes the total kinetic energy, which is described by $\frac{1}{2}M_{tot}v^2$; and U the total potential energy, which is described by $-\frac{1}{2}\frac{GM_{tot}^2}{R_{tot}}$. For a full derivation of the virial theorem see Collins (1978). Then, it follows that:

$$2\left(\frac{1}{2}M_{tot}v^2\right) - \frac{1}{2}\frac{GM_{tot}^2}{R_{tot}} = 0$$

$$M_{tot}v^2 = \frac{1}{2}\frac{GM_{tot}^2}{R_{tot}}$$

$$M_{tot} = \frac{2R_{tot}v^2}{G}$$

Now, using the reported values of van Dokkum et al. (2018a) of a $7.6kpc$ radius and a velocity dispersion of $10.5\frac{km}{s}$ and converting the units to International System of Units (SI) base units:

$$M_{tot} = \frac{2(7.6kpc)(3.0857\times10^{19}\frac{m}{kpc})((10.5\frac{km}{s})(1,000\frac{m}{km}))^2}{6.6740\times10^{-11}\frac{m^3}{kgs^2}}$$

$$M_{tot} = \frac{2(2.3451\times10^{20}m)(10,500\frac{m}{s})^2}{6.6740\times10^{-11}\frac{m^3}{kgs^2}}$$

$$M_{tot} = \frac{5.1709\times10^{28}\frac{m^3}{s^2}}{6.6740\times10^{-11}\frac{m^3}{kgs^2}}$$

$$M_{tot} = 7.7479\times10^{38}kg$$

Or expressed in solar masses, $M\odot \approx 1.9884x10^{30}kg$:

$$M_{tot} = 3.8965\times10^{8}\odot$$

which is in the same order of magnitude as van Dokkum et al.'s (2018a) result of $3.4x10^{8}\odot$ using the TME method. This serves as a quick order of magnitude verification of their results. Lastly, the stellar-to-halo mass ratio (SHM), $\frac{M_{halo}}{M_{stars}}$, can be computed using the reported stellar mass value of $2x10^{8}\odot$:

$$\frac{M_{halo}}{M_{stars}} = \frac{3.8965\times10^{8}\odot}{2\times10^{8}\odot} \approx 1.95$$

This result shows that NGC 1052–DF2 is a significantly dark-matter-deficient galaxy since, for comparison, the Milky Way has an SHM ratio of approximately 20 (McMillan, 2011) and low-mass galaxies have an SHM ratio greater than 30 (Behroozi et al., 2013; Sardone et al., 2019).

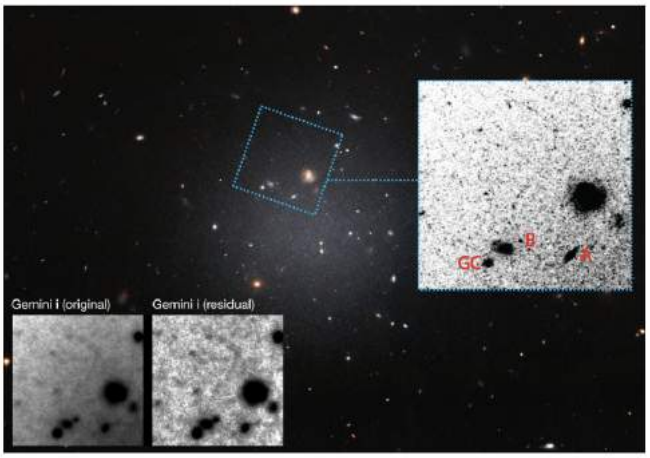

Figure 7. Image of NGC 1052-DF2's cold stellar stream that forms an inverted S-shape. Source: Abraham et al., 2018.

Besides its deficiency in non-baryonic matter, the galaxy NGC 1052–DF2 exhibits other unusual characteristics. Firstly, Abraham et al.

(2018) suggested that this galaxy possibly contains a thin cold stellar stream (Figure 6), which is the remnant of a dwarf galaxy or globular cluster pulled apart by tidal forces. If confirmed, it would replace the Pisces Stellar Stream (PSS), located at a distance of $35kpc$ (Martin et al., 2013), as the furthest cold stellar stream discovered. Secondly, its population of globular clusters (GCs) is very bright with an absolute magnitude of ≈ -9.1 (van Dokkum et al., 2018b), different from the expected value of -7.5 (Rejkuba, 2012). Since the 1970s, empirical evidence had lead to the general belief that the magnitude of old globular clusters peaked at M~ -7.5 (Rejkuba, 2012) across different types of galaxies. This became known as the so-called Globular Cluster Luminosity Function and has been used as a secondary distance indicator. Hence, NGC 1052–DF2 demonstrates that the luminosity function of metal-poor GCs is not a general law. Lastly, Sardone et al. (2019) found that the galaxy is extremely gas-poor, with a value of less than 1% of neutral atomic hydrogen, which is comparable to gas-poor elliptical dwarf galaxies. A low neutral hydrogen content has been linked to dense and crowded environments (Jung et al., 2018), thus revealing hints of NGC 1052–DF2's surroundings.

However, most of these irregularities would disappear if the UDG was located significantly closer to Earth than the 20 Mpc initially reported. Trujillo et al. (2019) argued that van Dokkum et al. (2018a) used the surface brightness fluctuation (SBF) technique (see Blakeslee (2012) for a review) to calculate the distance in a calibration range where its applicability is untested. Instead, Trujillo's team used redshift-independent techniques and concluded that NGC 1052–DF2

is located at 13 Mpc. In response, van Dokkum et al. (2018c) explained that Trujillo et al. (2019) used a colour-magnitude diagram giving erroneous distances due to blends. Using another distance ladder method, van Dokkum et al. (2018c) obtained a value of 18.7 Mpc. Furthermore, an independent study by Blakeslee & Cantiello (2018) agreed with the original distance. It seems that this original value is the correct one since, as of late, Shen et al. (2021) further agreed with these findings when they reported a distance of 22 Mpc with the tip of the red giant branch (TRGB) technique (see Salaris & Cassisi (1997) for an introduction) using the Hubble Space Telescope.

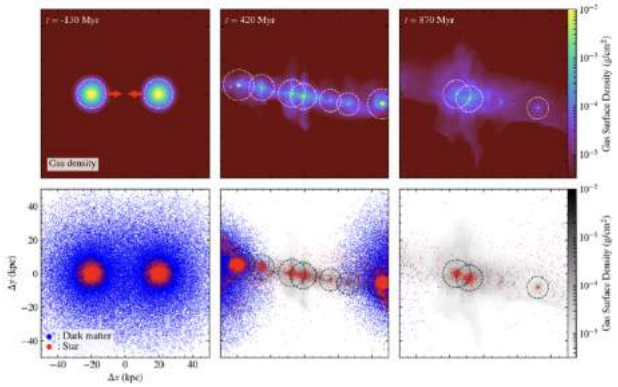

Figure 8. Galaxy collision simulation based on gravito-hydrodynamics using the ENZO code. Both panels show the gas surface density plotted against location, but the bottom panel distinguishes dark matter (in blue) from baryonic matter (in red). In the last box, dark matter deficient galaxies formed. Source: Shin et al., 2020.

While typical UDGs might form either through the failed giant galaxy scenario (van Dokkum et al., 2015) or the extended dwarf galaxy scenario (Amorisco & Loeb, 2016), as discussed in the previous section, NGC 1052–DF2 requires additional mechanisms to explain its lack of dark matter. The first possibility is that the galaxy lost its non-baryonic content through tidal stripping (Ogiya, 2018), which occurs when a galaxy tears apart material from a less massive one due to hefty tidal forces. Another possibility is that the UDG initially formed as a dwarf galaxy in a gas cloud, expelled due to energetic outflows from a quasar (Natarajan et al., 1998) or tidal interactions (Duc et al., 2014). Lastly, NGC 1052–DF2 could also have formed through the collision of gas-rich galaxies at high velocities (Silk, 2019). As Figure 8 shows, Shin et al. (2020) communicated that—using the ENZO and GADGET-2 codes—dark matter deficient galaxies with masses up to $10^8 \odot$ can form if two gas-rich galaxies collide at a velocity of $300 \frac{km}{s}$. Hence, there are several methods to remove the dark matter contents of a UDG, and they might not be mutually exclusive. Furthermore, while all the previous formation mechanisms rely on dynamical interactions (i.e., collisions, mergers, expulsions) to explain the deficiency of dark matter, non-baryonic matter is essential for UDG formation, albeit removed at a later stage.

Having discussed the characteristics of the NGC 1052–DF2 galaxy, I will now compare and contrast the predictions of CDM and MOND for this system. For a review of the basic tenets of each paradigm, I refer the reader to sections 1 and 2.

Within the CDM model, dark matter deficient galaxies are theorised in the dual dwarf galaxy theorem (Kroupa, 2012). This theorem predicts two different types of dwarf objects: primordial and tidal dwarf galaxies (TDGs). Primordial galaxies are formed via the usual clustering process of CDM, where such particles collapse into halos and those, in turn, trap baryonic matter. Hence, these types of structures are dark matter-dominated. In comparison, TDGs form through the dynamical interactions of gas-rich massive galaxies, causing powerful tidal forces that eject stellar and gaseous content. Haslbauer et al. (2019a) explained that this expelled material forms "tidal tails and arms" (p. 1), which continues to orbit its host galaxy and grow in mass, eventually becoming a dwarf. Two key characteristics of these TDGs are their dark matter-deficiency and low metallicity values if formed long ago ($\sim > 6Gyr$ ago) (Hunter et al., 2000). NGC 1052–DF2 meets the criteria with an estimated age of $8.9Gyr$ and a low metallicity of $[M/H] = -1.63$ (Fensch et al., 2019). Hence, I argue that this UDG, which is probably dwarf-scale sized (Chamba et al., 2019), could be considered a TDG. Consequently, the existence of the UDG NGC 1052–DF2 is in agreement with the CDM model.

The results of N-body simulations, nonetheless, show mixed results. Haslbauer et al. (2019a) found dark matter-deficient galaxies in the Illustris simulation, albeit in only 0.17% of all galaxies. Similarly, Jing et al. (2019) found only 2% of such objects in the EAGLE and Illustris simulations. However, in a follow-up paper, Haslbauer et al. (2019b) claimed that, in fact, in the TNG100-1 simulation, the probability of finding a UDG like NGC 1052–DF2 is much lower—a mere 4.8×10^{-7}. In contrast, Shin et al.'s (2020) collision-induced simulations successfully found dark matter-deficient galaxies with the ENZO and GADGET-2 codes. Intriguingly, given a dense initial environment, Macciò et al.'s (2020) high-resolution cosmological hydrodynamical simulations, enhanced with direct interactions, were able to reproduce "all key properties" (p. 699) of NGC 1052–DF2.

Despite the ambivalent results of simulations, NGC 1052–DF2 is in good agreement with the theoretical principles of the CDM model. Furthermore, the announcement that a second UDG lacks dark matter (van Dokkum et al., 2019) strengthens the idea that these galaxies probably constitute a new class of objects, namely TDGs, and are not merely exceptions, as they were previously considered.

I will now discuss NGC 1052–DF2 in the context of a MONDian universe. Notably, primordial galaxies and TDGs are now undistinguishable since baryonic matter alone determines the characteristics of a galaxy. The lack of dark matter in this scenario means that the formation and development of galaxies depend only on baryonic matter. Consequently, primordial galaxies—which are DM-dominated—and TDGs—which are DM-deficient—are undistinguishable for the MONDian theory. Renaud et al. (2016) argued that

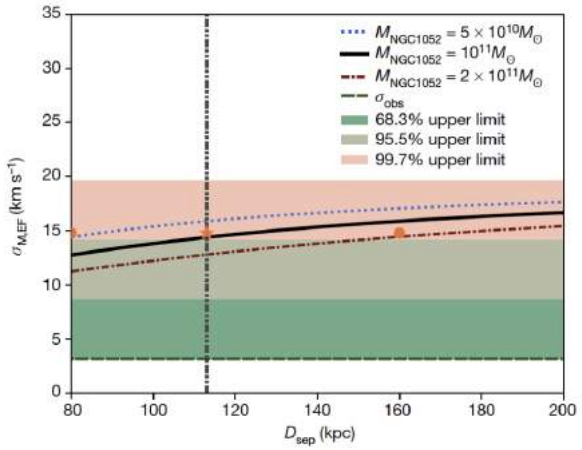

Figure 9. The MONDian velocity dispersion assumes the external field effect (EFE) to be a function of the separation between NGC 1052–DF and its elliptical host NGC 1052. The orange dots represent the predicted velocity dispersion values for a separation of 80 and 160 kpc, respectively. The star represents the velocity dispersion value for the distance of 113 kpc, considered a more probable separation. Source: Kroupa et al., 2018.

TDGs arise more naturally in a MONDian context since there is more substructure created in tidal tides, and they are more resistant against the tidal forces of their hosts. Furthermore, Haslbauer et al. (2019a) claimed that TDGs arise less naturally in the CDM model because they are more likely to be destroyed in dynamical interactions, like mergers, due to the dynamical friction of dark matter halos.

Another point worth clarifying is van Dokkum et al.'s (2018a) premature calculation of a velocity dispersion that falsified MOND. As Haghi et al. (2019) rightfully pointed out, the external field effect (EFE) should always be included in the calculation of the velocity dispersion when dealing with a non-isolated body. This is the case for NGC 1052–DF2 as it is 80 kpc away from its massive host NGC 1052. Accordingly, Haghi et al. (2019) arrived at a theoretical velocity dispersion value of $14\frac{km}{s}$, thoroughly consistent with the observations, as shown in Figure 9.

However, it is still possible that NGC 1052–DF2 is isolated after all, in which case its separation from any massive galaxy is $> 300kpc$. Shen et al.'s (2021) recent findings demonstrate that NGC 1052–DF2 and NGC 1052–DF4 are 2.1 Mpc $(2100kpc)$ apart. This separation implies that one of these UDGs is probably not bound to the group since the virial diameter—the diameter up to which the virial theorem is valid—of NGC 1052 is $780kpc$. Hence, either NGC1052–DF2 or NGC1052–DF4 is bound to the system, but not both, as this would require a larger virial diameter than the one observed (see Figure 10). Forbes et al. (2019) suggested the possibility that NGC 1052–DF2 is unbound due to its high recession velocity, and noted that the real 3D recession velocity value could be even $\frac{3}{2}^{\frac{1}{2}}$ times larger. If NGC 1052-DF2 is indeed isolated, its predicted velocity dispersion can be calculated through the following expression (Haghi et al., 2019):

$$\sigma_M = \left(\frac{4}{81} GMa_0\right)^{\frac{1}{4}} (1 + 0.56exp(3.02x))^{0.184}$$
$$x \equiv \log_{10}(\frac{a_i}{a_0})$$

where G is Newton's gravitational constant; M is the total mass of the system; a_i is the internal acceleration of the object; and a_0 refers to the MONDian fundamental acceleration constant. Taking the reported values of $a_0 = 1.21 \times 10^{-10} \frac{m}{s^2}$ and $a_i = 0.15a_0$ from Kroupa et al. (2018), it follows that:

$$x = log_{10}\left(\frac{0.15a_0}{a_0}\right) = -0.8239$$

$$\sigma_M = (\frac{4}{81}(6.6740 \times 10^{-11} \frac{m^3}{kgs^2})(6.76 \times 10^{38} kg)$$
$$(1.21 \times 10^{-10} \frac{m}{s^2})^{\frac{1}{4}}(1 + 0.56exp(3.02(-0.82339)))^{0.184}$$

$$\sigma_M = \left(2.6958 \times 10^{17} \frac{m^4}{s^4}\right)^{\frac{1}{4}} (1 + 0.0465)^{0.184}$$

$$\sigma_M = \left(2.6958 \times 10^{17} \frac{m^4}{s^4}\right)^{\frac{1}{4}} (1 + 0.0465)^{0.184}$$

$$\sigma_M = 22,786.20 \frac{m}{s}(1.0465)^{0.184}$$

$$\sigma_M = 22,786.20 \frac{m}{s}(1.00840) = 22,977.6 \frac{m}{s} = 22.97 \frac{km}{s}$$

We then arrive at a value for σ_M of $22.97\frac{km}{s}$, which was van Dokkum et al.'s (2018a) original conclusion. Nevertheless, as Figure 9 depicts, Shen et al. (2021) were careful to contemplate a scenario where NGC1052–DF2 is under the influence of NGC 1052, and NGC1052–DF4 under that of NGC 1035, as has been proposed in the literature (Montes et al., 2020). There are, nonetheless, difficulties with this scenario since it would mean that the two UDGs formed independently, despite their "near-identical morphologies, kinematics, and extreme globular cluster population" (Shen et al., 2021, p. 10). Still, unless either NGC1052–DF2 or NGC1052–DF4 are proven to be isolated, MOND agrees with observations.

5 Summary and Conclusions

The study of dark-matter-deficient galaxies is essential to understand galaxy formation and the true nature of non-baryonic matter. This thesis introduced the reader to the recent debate that arose following the discovery that an ultra-diffuse galaxy (UDG) lacks dark matter, and is the first systematic literature review that focuses on the galaxy NGC 1052–DF2. This galaxy exhibits remarkable characteristics: for example, it is extremely gas-poor, it harbours a population of unusually bright globular clusters, and it might contain the most distant cold stellar stream known so far. Furthermore, it has important implications within the Cold Dark Matter (CDM) and Modified Newtonian dynamics (MOND) paradigms.

This paper concluded that, currently, both paradigms are in agreement with observations. On the one hand, CDM's

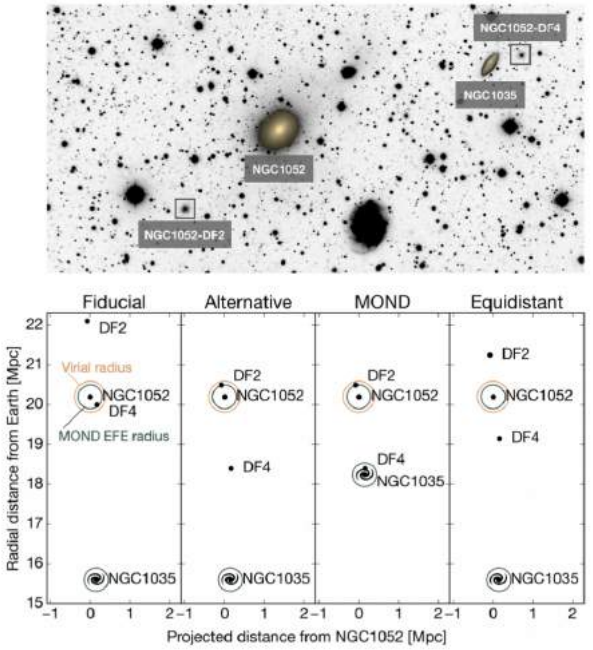

Figure 10. The top panel is an image of the NGC 1052 group taken with Dragonfly Telephoto Array. The bottom panel summarises the possible configurations of the two UDGs concerning the more massive galaxies NGC 1035 and NGC 1052. Only in one configuration is MOND consistent with observations (third arrangement). Source: Shen et al., 2021.

dual dwarf galaxy theorem predicts the existence of primordial galaxies and TDGs. Since the accurate size of NGC 1052–DF2 is probably in the dwarf galaxy scale, and because it meets the key characteristics of TDGs, it was argued that it should be considered a TDG. Hence, the existence of NGC 1052–DF2 proves a crucial aspect of the theory. Despite the initially mixed results of N-body simulations, the discovery of NGC 1052–DF4 (the second UDG found to be dark-matter-deficient) and more favourable simulation results increasingly support the idea that NGC 1052–DF2 represents a new class of objects and is not just an outlier. On the other hand, MOND's velocity dispersion predictions for NGC 1052–DF2 are correct once the external field effect (EFE) is considered. In fact, MOND has been exceedingly successful in explaining the dynamics of dwarf galaxies, and TDGs should arise more naturally in this framework.

The debate, however, continues considering the recent finding that NGC 1052–DF2 and NGC 1052–DF4 are separated by a few Mpc—a distance significantly larger than the virial diameter of their host. This finding implies that one of them is probably not bound to the 1052 group, in which case MOND's velocity dispersion prediction for the isolated object would not match observations. In comparison, CDM's predictions are unaffected by this observation, as they do not depend on whether these galaxies are isolated or not. However, more research is needed to determine beyond any reasonable doubt if either of the two UDGs are indeed isolated from the influence of any other massive galaxy and, if so, which one is isolated. In this sense, the upcoming generation of astronomical telescopes, like the James Webb, will shed light on this issue by elucidating the distance of NGC 1052–DF2 and NGC 1052–DF4 to nearby galaxies, thus advancing the modified gravity debate.

Acknowledgments

To begin with, I would like to thank my supervisor, Dr. Gianfranco Bertone, for introducing me to the thrilling world of dark matter. This thesis project could not have come to fruition without his valuable comments and guidance.

To all my lecturers, including my tutor Dr. Bart Verheggen, thank you for giving me the knowledge and tools needed to complete my undergraduate degree.

I am very grateful to the ASF AUC Scholarship Fund for funding my studies and allowing me to be part of the genuinely enriching and close-knit community of AUC.

I would also like to thank my family and friends who encouraged me to follow my wild dream of changing careers and studying physics abroad. To my brother and sister, thank you for your companionship that transcends borders. To my parents, I owe you all my successes, and I am only here thanks to your unconditional support. Finally, to William, thank you for your relentless encouragement and for always believing in me.

References

Abraham, R., Danieli, S., Dokkum, P. V., Conroy, C., Kruijssen, J. M., Cohen, Y., Merritt, A., Zhang, J., Lokhorst, D., Mowla, L., Brodie, J., Romanowsky, A. J., & Janssens, S. (2018). The maybe stream: A possible cold stellar stream in the ultra-diffuse galaxy NGC1052-DF2. *Research Notes of the AAS, 2*(2), 16. https://doi.org/10.3847/2515-5172/aac087

Abramoff, O., Barak, L., Bloch, I. M., Chaplinsky, L., Crisler, M., Dawa, Drlica-Wagner, A., Essig, R., Estrada, J., Etzion, E., Fernandez, G., Gift, D., Sofo-Haro, M., Taenzer, J., Tiffenberg, J., Volansky, T., & Yu, T. (2019). SENSEI: Direct-detection constraints on sub-gev dark matter from a shallow underground run using a prototype Skipper CCD. Physical *Review Letters, 122*(16). https://doi.org/10.1103/physrevlett.122.161801

Amorisco, N. C., & Loeb, A. (2016). Ultradiffuse galaxies: The high-spin tail of the abundant dwarf galaxy population. Monthly Notices of the Royal Astronomical Society: *Letters, 459*(1), L51-L55. https://doi.org/10.1093/mnrasl/slw055

Aprile, E., Aalbers, J., Agostini, F., Alfonsi, M., Althueser, L., Amaro, F., Antochi, V., Angelino, E., Angevaare, J., Arneodo, F., Barge, D., Baudis, L., Bauermeister, B., Bellagamba, L., Benabderrahmane, M., Berger, T., Brown,

A., Brown, E., Bruenner, S., ... Zopounidis, J. (2020). Projected WIMP sensitivity of the XENONnT dark matter experiment. *Journal of Cosmology and Astroparticle Physics, 2020*(11), 031-031. https://doi.org/10.1088/1475-7516/2020/11/031

Babcock, H. W. (1939). The rotation of the Andromeda nebula. Lick Observatory Bulletins, 19, 41-51. https://doi.org/10.5479/ads/bib/1939licob.19.41b

Bahcall, N. A. (2015). Dark matter universe. *Proceedings of the National Academy of Sciences of the United States of America, 112*(40), 12243-12245.

Behroozi, P. S., Wechsler, R. H., & Conroy, C. (2013). The average star formation histories of galaxies in dark matter halos FROMz= 0-8. *The Astrophysical Journal, 770*(1), 57. https://doi.org/10.1088/0004-637x/770/1/57

Bekenstein, J. D. (2004). Relativistic gravitation theory for the modified newtonian dynamics paradigm. *Physical Review D, 70*(8). https://doi.org/10.1103/physrevd.70.083509

Bertone, G., & Hooper, D. (2018). History of dark matter. *Reviews of Modern Physics, 90*(4), 045002. https://doi.org/10.1103/RevModPhys.90.045002

Blakeslee, J. P. (2012). Surface brightness fluctuations as primary and secondary distance indicators. *Astrophysics and Space Science, 341*(1), 179-186. https://doi.org/10.1007/s10509-012-0997-6

Blakeslee, J. P., & Cantiello, M. (2018). Independent analysis of the distance to NGC 1052-DF2. *Research Notes of the AAS, 2*(3), 146. https://doi.org/10.3847/2515-5172/aad90e

Blumenthal, G. R., Pagels, H., & Primack, J. R. (1982). Galaxy formation by dissipationless particles heavier than neutrinos. *Nature, 299*(5878), 37-38. https://doi.org/10.1038/299037a0

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. (1984). Formation of galaxies and large-scale structure with cold dark matter. *Nature, 311*(5986), 517-525. https://doi.org/10.1038/311517a0

Bond, J. R., Szalay, A. S., & Turner, M. S. (1982). Formation of galaxies in a gravitino-dominated universe. *Physical Review Letters, 48*(23), 1636-1639. https://doi.org/10.1103/physrevlett.48.1636

Burles, S., Nollett, K. M., & Turner, M. S. (2001). Big Bang nucleosynthesis predictions for precision cosmology. *The Astrophysical Journal, 552*(1), L1-L5. https://doi.org/10.1086/320251

Chaboyer, B., Demarque, P., Kernan, P. J., & Krauss, L. M. (1998). The age of globular clusters in light ofHipparcos: Resolving the age problem? *The Astrophysical Journal, 494*(1), 96-110. https://doi.org/10.1086/305201

Chamba, N., Trujillo, I., & Knapen, J. H. (2020). Are ultra-diffuse galaxies milky way-sized? *Astronomy & Astrophysics*, 633, L3. https://doi.org/10.1051/0004-6361/201936821

Clowe, D., Gonzalez, A., & Markevitch, M. (2004). Weak-Lensing mass reconstruction of the interacting cluster 1E 0657−558: Direct evidence for the existence of dark matter. *The Astrophysical Journal, 604*(2), 596-603. https://doi.org/10.1086/381970

Collins, G. W. (1978). The Virial theorem in stellar astrophysics. *Pachart Publishing House*.

Crisosto, N., Sikivie, P., Sullivan, N., Tanner, D., Yang, J., & Rybka, G. (2020). ADMX SLIC: Results from a superconducting LC circuit investigating cold Axions. *Physical Review Letters, 124*(24). https://doi.org/10.1103/physrevlett.124.241101

Cui, W., Knebe, A., Yepes, G., Yang, X., Borgani, S., Kang, X., Power, C., & Staveley-Smith, L. (2017). The large-scale environment from cosmological simulations – I. The baryonic cosmic web. *Monthly Notices of the Royal Astronomical Society, 473*(1), 68-79. https://doi.org/10.1093/mnras/stx2323

Di Cintio, A., Brook, C., Dutton, A. A., Macciò, A. V., Obreja, A., & Dekel, A. (2017). NIHAO – XI. Formation of ultra-diffuse galaxies by outflows. *Monthly Notices of the Royal Astronomical Society: Letters, 466*(1), L1-L6. https://doi.org/10.1093/mnrasl/slw210

Duc, P., Paudel, S., McDermid, R. M., Cuillandre, J., Serra, P., Bournaud, F., Cappellari, M., & Emsellem, E. (2014). Identification of old tidal dwarfs near early-type galaxies from deep imaging and H i observations. *Monthly Notices of the Royal Astronomical Society, 440*(2), 1458-1469. https://doi.org/10.1093/mnras/stu330

Famaey, B., & McGaugh, S. S. (2012). Modified newtonian dynamics (MOND): Observational phenomenology and relativistic extensions. *Living Reviews in Relativity, 15*(1). https://doi.org/10.12942/lrr-2012-10

Feng, J. L. (2010). Dark matter candidates from particle physics and methods of detection. *Annual Review of Astronomy and Astrophysics, 48*(1), 495-545. https://doi.org/10.1146/annurev-astro-082708-101659

Fensch, J., Van der Burg, R. F., Jeřábková, T., Emsellem, E., Zanella, A., Agnello, A., Hilker, M., Müller, O., Rejkuba, M., Duc, P., Durrell, P., Habas, R., Lim, S., Marleau, F. R., Peng, E. W., & Sánchez Janssen, R. (2019). The ultra-diffuse galaxy NGC 1052-DF2 with MUSE. *Astronomy & Astrophysics*, 625, A77. https://doi.org/10.1051/0004-6361/201834911

Forbes, D., Alabi, A., Brodie, J., & Romanowsky, A. (2019). Dark matter and no dark matter: on the halo mass of NGC 1052. *Monthly Notices of the Royal Astronomical Society, 489*(3), 3665-3669. https://doi.org/10.1093/mnras/stz2420

Gentile, G., Famaey, B., & De Blok, W. J. (2011). THINGS about MOND. *Astronomy & Astrophysics*, 527, A76. https://doi.org/10.1051/0004-6361/201015283

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D, 23*(2), 347-356. https://doi.org/10.1103/physrevd.23.347

Gunn, J. E. (1980). The dynamics of galaxies and the 'missing mass' problem. Philosophical Transactions of the Royal Society of London. *Series A, Mathematical and Physical Sciences, 296*(1419), 313-318. https://doi.org/10.1098/rsta.1980.0175

Haghi, H., Kroupa, P., Banik, I., Wu, X., Zonoozi, A. H., Javanmardi, B., Ghari, A., Müller, O., Dabringhausen, J., & Zhao, H. (2019). A new formulation of the external field effect in mond and numerical simulations of ultra-diffuse dwarf galaxies – application to NGC 1052-df2 and NGC 1052-df4. *Monthly Notices of the Royal Astronomical Society, 487*(2), 2441-2454. https://doi.org/10.1093/mnras/stz1465

Haslbauer, M., Dabringhausen, J., Kroupa, P., Javanmardi, B., & Banik, I. (2019a). Galaxies lacking dark matter in the Illustris simulation. *Astronomy & Astrophysics*, 626, A47. https://doi.org/10.1051/0004-6361/201833771

Haslbauer, M., Banik, I., Kroupa, P., & Grishunin, K. (2019b). The ultra-diffuse dwarf galaxies NGC 1052-DF2 and 1052-DF4 are in conflict with standard cosmology. *Monthly Notices of the Royal Astronomical Society, 489*(2), 2634-2651. https://doi.org/10.1093/mnras/stz2270

Hunter, D. A., Hunsberger, S. D., & Roye, E. W. (2000). Identifying old tidal dwarf irregulars. *The Astrophysical Journal, 542*(1), 137-142. https://doi.org/10.1086/309542

Jing, Y., Wang, C., Liao, S., Wang, J., Guo, Q., & Gao, L. (2019). Dark-matter-deficient galaxies in hydrodynamical simulations. *Monthly Notices of the Royal Astronomical Society, 488*(3), 3298-3307. https://doi.org/10.1093/mnras/stz1839

Jung, S. L., Choi, H., Wong, O. I., Kimm, T., Chung, A., & Yi, S. K. (2018). On the origin of gas-poor galaxies in galaxy clusters using cosmological hydrodynamic simulations. *The Astrophysical Journal, 865*(2), 156. https://doi.org/10.3847/1538-4357/aadda2

Kaviraj, S. (2020). The low-surface-brightness Universe: a new frontier in the study of galaxy evolution. *arXiv*. https://arxiv.org/pdf/2001.01728.pdf

Kim, S. Y., Peter, A. H., & Hargis, J. R. (2018). Missing satellites problem: Completeness corrections to the number of satellite galaxies in the Milky Way are consistent with cold dark matter predictions. *Physical Review Letters, 121*(21). https://doi.org/10.1103/physrevlett.121.211302

Kroupa, P. (2012). The dark matter crisis: Falsification of the current Standard Model of cosmology. *Publications of the Astronomical Society of Australia, 29*(4), 395-433. https://doi.org/10.1071/as12005

Kroupa, P., Haghi, H., Javanmardi, B., Zonoozi, A. H., Müller, O., Banik, I., Wu, X., Zhao, H., & Dabringhausen, J. (2018). Does the galaxy NGC1052–DF2 falsify Milgromian dynamics? *Nature, 561*(7722), E4-E5. https://doi.org/10.1038/s41586-018-0429-z

Lawrence, A., Warren, S. J., Almaini, O., Edge, A. C., Hambly, N. C., Jameson, R. F., Lucas, P., Casali, M., Adamson, A., Dye, S., Emerson, J. P., Foucaud, S., Hewett, P., Hirst, P., Hodgkin, S. T., Irwin, M. J., Lodieu, N., McMahon, R. G., Simpson, C., ... Folger, M. (2007). The UKIRT infrared deep sky survey (UKIDSS). *Monthly Notices of the Royal Astronomical Society, 379*(4), 1599-1617. https://doi.org/10.1111/j.1365-2966.2007.12040.x

Lim, S., Côté, P., Peng, E. W., Ferrarese, L., Roediger, J. C., Durrell, P. R., Mihos, J. C., Wang, K., Gwyn, S. D., Cuillandre, J., Liu, C., Sánchez-Janssen, R., Toloba, E., Sales, L. V., Guhathakurta, P., Lançon, A., & Puzia, T. H. (2020). The next generation Virgo cluster survey (NGVS). XXX. ultra-diffuse galaxies and their globular cluster systems. *The Astrophysical Journal, 899*(1), 69. https://doi.org/10.3847/1538-4357/aba433

Lisanti, M., Moschella, M., Outmezguine, N. J., & Slone, O. (2019). Testing dark matter and modifications to gravity using local Milky Way observables. *Physical Review D, 100*(8). https://doi.org/10.1103/physrevd.100.083009

Ma, H., She, Z., Liu, Z., Yang, L., Yue, Q., Zeng, Z., & Xue, T. (2020). Results of direct dark matter detection with CDEX experiment at CJPL. *Journal of Physics: Conference Series*, 1468, 012070. https://doi.org/10.1088/1742-6596/1468/1/012070

Macciò, A. V., Prats, D. H., Dixon, K. L., Buck, T., Waterval, S., Arora, N., Courteau, S., & Kang, X. (2020). Creating a galaxy lacking dark matter in a dark matter-dominated universe. *Monthly Notices of the Royal Astronomical Society, 501*(1), 693-700. https://doi.org/10.1093/mnras/staa3716

Mancera Piña, P. E., Peletier, R. F., Aguerri, J. A., Venhola, A., Trager, S., & Choque Challapa, N. (2018). Reviewing the frequency and central depletion of ultra-diffuse galaxies in galaxy clusters from the KIWICS survey. *Monthly Notices of the Royal Astronomical Society, 481*(4), 4381–4388. https://doi.org/10.1093/mnras/sty2574

Martin, C., Carlin, J. L., Newberg, H. J., & Grillmair, C. (2013). Kinematic discovery of a stellar stream located in Pisces. *The Astrophysical Journal, 765*(2), L39. https://doi.org/10.1088/2041-8205/765/2/l39

Martin, G., Kaviraj, S., Laigle, C., Devriendt, J. E., Jackson, R. A., Peirani, S., Dubois, Y., Pichon, C., & Slyz, A. (2019). The formation and evolution of low-surface-brightness galaxies. *Monthly Notices of the Royal Astronomical Society, 485*(1), 796–818. https://doi.org/10.1093/mnras/stz356

McGaugh, S. (2020). Predictions and outcomes for the dynamics of rotating galaxies. *Galaxies, 8*(2), 35. https://doi.org/10.3390/galaxies8020035

McMillan, P. J. (2011). Mass models of the Milky Way. *Monthly Notices of the Royal Astronomical Society, 414*(3), 2446-2457. https://doi.org/10.1111/j.1365-2966.2011.18564.x

Milgrom, M. (1983). A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal*, 270, 365. https://doi.org/10.1086/161130

Milgrom, M. (2007). Milgrom's perspective on the Bullet Cluster. *The MOND Pages*. https://astroweb.case.edu/ssm/mond/moti_bullet.html

Milgrom, M. (2014). MOND theory. *Canadian Journal of Physics, 93*(2), 107-118. https://doi.org/10.1139/cjp-2014-0211

Montes, M., Infante-Sainz, R., Madrigal-Aguado, A., Román, J., Monelli, M., Borlaff, A. S., & Trujillo, I. (2020). The galaxy "Missing dark matter" NGC 1052-DF4 is undergoing tidal disruption. *The Astrophysical Journal, 904*(2), 114. https://doi.org/10.3847/1538-4357/abc340

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. (1999). Dark matter substructure within galactic halos. *The Astrophysical Journal, 524*(1), L19-L22. https://doi.org/10.1086/312287

Natarajan, P., Sigurdsson, S., & Silk, J. (1998). Quasar outflows and the formation of dwarf galaxies. *Monthly Notices of the Royal Astronomical Society, 298*(2), 577-582. https://doi.org/10.1046/j.1365-8711.1998.01703.x

National Aeronautics and Space Administration. (2010, April 16). *WMAP inflation theory*. `https://wmap.gsfc.nasa.gov/universe/bb_cosmo_infl.html`

Navarro, J. F., Frenk, C. S., & White, S. D. (1997). A universal density profile from hierarchical clustering. *The Astrophysical Journal, 490*(2), 493-508. https://doi.org/10.1086/304888

Netterfield, C. B., Ade, P. A., Bock, J. J., Bond, J. R., Borril, J., Boscaleri, A., Coble, K., Contaldi, C. R., Crill, B. P., & de Bernardis, P. (2002). A Measurement by BOOMERANG of Multiple Peaks in the Angular Power Spectrum of the Cosmic Microwave Background. *The Astrophysical Journal, 571*(2), 604. https://iopscience.iop.org/article/10.1086/340118

Ogiya, G. (2018). Tidal stripping as a possible origin of the ultra diffuse galaxy lacking dark matter. *Monthly Notices of the Royal Astronomical Society: Letters, 480*(1), L106-L110. https://doi.org/10.1093/mnrasl/sly138

Oh, S., Hunter, D. A., Brinks, E., Elmegreen, B. G., Schruba, A., Walter, F., Rupen, M. P., Young, L. M., Simpson, C. E., Johnson, M. C., Herrmann, K. A., Ficut-Vicas, D., Cigan, P., Heesen, V., Ashley, T., & Zhang, H. (2015). High-resolution mass models of dwarf galaxies from little things. *The Astronomical Journal, 149*(6), 180. https://doi.org/10.1088/0004-6256/149/6/180

Oman, K. A., Navarro, J. F., Fattahi, A., Frenk, C. S., Sawala, T., White, S. D., Bower, R., Crain, R. A., Furlong, M., Schaller, M., Schaye, J., & Theuns, T. (2015). The unexpected diversity of dwarf galaxy rotation curves. *Monthly Notices of the Royal Astronomical Society, 452*(4), 3650-3665. https://doi.org/10.1093/mnras/stv1504

Paraficz, D., Kneib, J., Richard, J., Morandi, A., Limousin, M., Jullo, E., & Martinez, J. (2016). The bullet cluster at its best: Weighing stars, gas, and dark matter. *Astronomy & Astrophysics*, 594, A121. https://doi.org/10.1051/0004-6361/201527959

Pearson, D. W., & Batuski, D. J. (2013). Locating bound structure in an accelerating universe. *Monthly Notices of the Royal Astronomical Society, 436*(1), 796-806. https://doi.org/10.1093/mnras/stt1614

Peebles, P. J. (1982). Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *The Astrophysical Journal*, 263, L1. https://doi.org/10.1086/183911

Planck Collaboration. (2018, July). [Map]. *European Space Agency*. `https://www.cosmos.esa.int/documents/387566/1753103/-Planck_2018_T_CMB.pdf/bd1c9c9d-db10-1396-61c0-902a7dd63d7b`

Primack, J. R., & Gross, M. S. (2001). Hot Dark Matter in Cosmology. *In Current Aspects of Neutrino Physics* (pp. 287-308). Springer. https://doi.org/10.1007/978-3-662-04597-8_12

Rejkuba, M. (2012). Globular cluster luminosity function as distance indicator. *Astrophysics and Space Science*, 341, 195–206. https://doi.org/10.1007/s10509-012-0986-9

Renaud, F., Famaey, B., & Kroupa, P. (2016). Star formation triggered by galaxy interactions in modified gravity.

Monthly Notices of the Royal Astronomical Society, 463(4), 3637-3652. https://doi.org/10.1093/mnras/stw2331

Román, J., & Trujillo, I. (2017). Ultra-diffuse galaxies outside clusters: Clues to their formation and evolution. *Monthly Notices of the Royal Astronomical Society, 468*(4), 4039-4047. https://doi.org/10.1093/mnras/stx694

Rubin, V. C., & Ford, W. K. (1970). Rotation of the Andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159, 379. https://doi.org/10.1086/150317

Salaris, M., & Cassisi, S. (1997). The 'tip' of the red giant branch as a distance indicator: Results from evolutionary models. *Monthly Notices of the Royal Astronomical Society, 289*(2), 406-414. https://doi.org/10.1093/mnras/289.2.406

Sanders, R. H., & Verheijen, M. A. (1998). Rotation curves of Ursa Major galaxies in the context of modified newtonian dynamics. *The Astrophysical Journal, 503*(1), 97-108. https://doi.org/10.1086/305986

Sanders, R. H., & McGaugh, S. (2002). An alternative to dark matter: Modified newtonian dynamics. *The Dark Matter Problem*, 40, 132-149. https://doi.org/10.1017/cbo9781139192309.010

Sardone, A., Pisano, D. J., Burke-Spolaor, S., Mascoop, J. L., & Pol, N. (2019). Constraints on the H I mass for NGC 1052-DF2. *The Astrophysical Journal, 871*(2), L31. https://doi.org/10.3847/2041-8213/ab0084

Shandarin, S., Habib, S., & Heitmann, K. (2010). Origin of the cosmic network inΛCDM: Nature vs nurture. *Physical Review D, 81*(10). https://doi.org/10.1103/physrevd.81.103006

Shen, Z., Danieli, S., Van Dokkum, P., Abraham, R., Brodie, J. P., Conroy, C., Dolphin, A. E., Romanowsky, A. J., Diederik Kruijssen, J. M., & Dutta Chowdhury, D. (2021). A tip of the red giant branch distance of 22.1 $\pm$ 1.2 Mpc to the dark matter deficient galaxy NGC 1052–DF2 from 40 orbits of Hubble space telescope imaging. *The Astrophysical Journal Letters, 914*(1), L12. https://doi.org/10.3847/2041-8213/ac0335

Shin, E., Jung, M., Kwon, G., Kim, J., Lee, J., Jo, Y., & Oh, B. K. (2020). Dark matter deficient galaxies produced via high-velocity galaxy collisions in high-resolution numerical simulations. *The Astrophysical Journal, 899*(1), 25. https://doi.org/10.3847/1538-4357/aba434

Silk, J. (2019). Ultra-diffuse galaxies without dark matter. *Monthly Notices of the Royal Astronomical Society: Letters, 488*(1), L24-L28. https://doi.org/10.1093/mnrasl/slz090

Skrutskie, M. F., Cutri, R. M., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J., Huchra, J., Liebert, J., Lonsdale, C., Monet, D. G., Price, S., Seitzer, P., Jarrett, T., Kirkpatrick, J. D., Gizis, J. E., Howard, E., ... Wheelock, S. (2006). The Two Micron All Sky Survey (2MASS). *The Astronomical Journal, 131*(2), 1163. https://iopscience.iop.org/article/10.1086/498708

Skordis, C., & Zlosnik, T. (2020). A new relativistic theory for Modified Newtonian Dynamics. *arXiv*. https://arxiv.org/abs/2007.00082

Smith, S. (1936). The mass of the virgo cluster. *The Astrophysical Journal*, 83, 23-30.

Tasitsiomi, A. (2003). The state of the cold dark matter models on galactic and subgalactic scales. *International Journal of Modern Physics D, 12*(07), 1157-1196. https://doi.org/10.1142/s0218271803003426

Trujillo, I., Beasley, M. A., Borlaff, A., Carrasco, E. R., Di Cintio, A., Filho, M., Monelli, M., Montes, M., Román, J., Ruiz-Lara, T., Sánchez Almeida, J., Valls-Gabaud, D., & Vazdekis, A. (2019). A distance of 13 Mpc resolves the claimed anomalies of the galaxy lacking dark matter. *Monthly Notices of the Royal Astronomical Society, 486*(1), 1192-1219. https://doi.org/10.1093/mnras/stz771

Tully, R. B., & Fisher, J. R. (1977). A new method of determining distances to galaxies. *Astronomy & Astrophysics, 54*(3), 661-673.

Van Dokkum, P. V., Abraham, R., Merritt, A., Zhang, J., Geha, M., & Conroy, C. (2015). Forty-seven Milky Way-sized, Extremely Diffuse Galaxies in the Coma Cluster. *The Astrophysical Journal Letters, 798*(2), L45. https://doi.org/10.1088/2041-8205/798/2/L45

Van Dokkum, P., Danieli, S., Cohen, Y., Merrit, A., Romanowsky, A. J., Abraham, R., Brodie, J., Conroy, C., Lokhorst, D., Mowla, L., O'Sullivan, E., & Zhang, J. (2018a). A galaxy lacking dark matter. *Nature*, 555, 629–632. https://doi.org/10.1038/nature25767

Van Dokkum, P., Cohen, Y., Danieli, S., Kruijssen, J. M., Romanowsky, A. J., Merritt, A., Abraham, R., Brodie, J., Conroy, C., Lokhorst, D., Mowla, L., O'Sullivan, E., & Zhang, J. (2018b). An enigmatic population of luminous globular clusters in a galaxy lacking dark matter. *The Astrophysical Journal, 856*(2), L30. https://doi.org/10.3847/2041-8213/aab60b

Van Dokkum, P., Danieli, S., Cohen, Y., Romanowsky, A. J., & Conroy, C. (2018c). The Distance of the Dark Matter Deficient Galaxy NGC 1052–DF2. *The Astrophysical Journal Letters, 864*(1), L18. https://doi.org/10.3847/2041-8213/aada4d

Van Dokkum, P., Danieli, S., Abraham, R., Conroy, C., & Romanowsky, A. J. (2019). A second galaxy missing dark matter in the NGC 1052 group. *The Astrophysical Journal,*

874(1), L5. https://doi.org/10.3847/2041-8213/ab0d92

Weinberg, D. H., Bullock, J. S., Governato, F., Kuzio de Naray, R., & Peter, A. H. (2015). Cold dark matter: Controversies on small scales. *Proceedings of the National Academy of Sciences, 112*(40), 12249-12255. https://doi.org/10.1073/pnas.1308716112

Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Phys. Acta*, 6, 110-127.

Social Sciences

Young Entrepreneurs of Reconciliation

The Role of Rwandan Youth in the Diffusion of Reconciliation Norms

Carolina Resigotti

Supervisor
Dr. Siniša Vuković (AUC)
Reader
Dr. Erica Pasini (AUC)

Photographer: Richard Essink

Abstract

This research expands on existing literature merging norm diffusion and reconciliation, by exploring the role of young people as norm entrepreneurs spreading norms of reconciliation. In order to do so, it examines the post-genocide context of Rwanda, and analyzes data drawn from semi-structured interviews with young Rwandans involved in the reconciliation process. The paper provides an overview of the globally recognized norms of reconciliation advanced by the Rwandan state, with the purpose of understanding the ways in which the new Rwandan generation engages in national and international norm promotion. Findings illustrate that youth actors localize norms of truth and memory, and internalize unity and a common identity. In line with global youth trends, they advance a stronger focus on a shared future marked by a "never again" narrative, and they also encourage empathy, critical thinking, and responsibility. By using innovative methods like art, dialogue, and media, the youth actors increasingly interact with the state's transitional justice institutions and the education system, as they stand between norm emergence and cascade. Their impact is also visible between cascade and internalization, as they form networks of norm entrepreneurs beyond national boundaries. Such results implicate the relevant agency of young people within the norm diffusion cycle, by virtue of their ability to reconstruct post-conflict societies. Lastly, findings contribute to emphasizing their essential role during reconciliation processes.

Keywords: *Norm diffusion, reconciliation, youth actors, norm entrepreneurs, Rwanda*

Suggested citation:
Resigotti, C. (2021). Young entrepreneurs of reconciliation: The role of Rwandan youth in the diffusion of reconciliation norms. *AUC Undergraduate Journal of Liberal Arts and Sciences, 16*(1), 46-68.

Contents

List of Figures and Tables

Abbreviations

CNLG	National Commission for the Fight against Genocide
DRC	Democratic Republic of the Congo
EU	European Union
ICC	International Criminal Court
ICTR	International Criminal Tribunal for Rwanda
ICTY	International Criminal Tribunal for the former Yugoslavia
NGO	Non-Governmental Organization
NURC	National Unity and Reconciliation Commission
PLP	Peace and Love Proclaimers
SD	Sustained Dialogue
TAN	Transnational Advocacy Network
TRC	Truth and Reconciliation Commission
UN	United Nations
UNESCO	United Nations Educations, Scientific and Cultural Organization
UNOY	United Network of Young Peacebuilders
UNSC	United Nations Security Council
UNSCR	United Nations Security Council Resolution
YPS	Youth, Peace and Security

1 Introduction

Reconciliation is an indispensable requirement for the reconstruction of society in the aftermath of a violent conflict. By virtue of its power to restore social relationships and modify societal beliefs, attitudes, motivations, and emotions (Bar-Tal & Bennink, 2004; Rosoux, 2009; 2015; 2017), reconciliation is internationally viewed as an essential condition for the attainment of stable peace (Rosoux, 2017). This global consensus around the significance of reconciliation hints at its association with norm diffusion.

The latter indeed relates to the emergence, propagation, and internalization of "standards of appropriate behavior" (Finnemore & Sikkink, 1998, p. 891) within the international realm. According to Finnemore and Sikkink's (1998) norm life cycle, norms are created by norm entrepreneurs,"cascade" across nation-states through the work of international organizations and networks and are internalized within national and local social discourses. In order to explain how norms can be dynamically negotiated between the global and the local, existing literature has analyzed processes of norm contestation (Stimmer, 2019; Wiener, 2004; Wiener, 2017), and norm localization (Acharya, 2004; Capie, 2008; Groß, 2015; Tholens & Groß, 2015).

The link between norm diffusion and reconciliation becomes clear through the acknowledgement of seemingly internationally agreed-upon norms of reconciliation, or standards for its effectiveness, such as justice (Bar-Tal & Bennink, 2004; Rosoux, 2009; Staub, 2013; Weinstein & Stover, 2004), truth and collective memory (Bar-Tal & Bennink, 2004; Kelman, 2004; Rosoux, 2009; Staub, 2013), and forgiveness (Auerbach, 2004; Bar-Tal & Bennink, 2004; Rosoux & Brudholm, 2009; Rosoux, 2009; 2015). These normative elements are reinforced through the process of transitional justice, which itself is globally considered a post-conflict norm (Ben-Josef Hirsch, 2014; Boesenecker & Vinjamuri, 2011; Kostovicova & Bicquelet, 2017; MacKenzie & Sesay, 2012; Ottendörfer, 2013; Sikkink & Kim, 2013; Subotic, 2015) Evidence for this global conception is offered by the establishment of international tribunals in order to ensure legal justice (Ben-Josef Hirsch, 2014; Sikkink & Kim, 2013), the institutionalization of truth and reconciliation commissions (TRCs) for the exposure of truth and the necessity to forgive (Ben-Josef Hirsch, 2014; Rosoux & Brudholm, 2009, p. 34), and the implementation of museums and memorial sites for the propagation of collective memory (Sikkink & Kim, 2013, p. 270).

Bearing the connection between reconciliation and norm diffusion in mind, this research aims at more thoroughly exploring the role of local actors, specifically the youth. Even though youth agency is often disregarded in the national and international spheres (De Graaf, 2014; McEvoy-Levy, 2006; Mengistu, 2017), their role has proven to be crucial in post-conflict areas, during peacebuilding and reconciliation processes (De Graaf, 2014; Dragija, 2020; Kasherwa, 2019; Kosic & Tauber, 2010; Kurze, 2016; McEvoy-Levy, 2006; Nguyen-Marshall, 2015; Prisca et al., 2012; Shipler, 2008; Thapa, 2009; Wienand, 2013; Wollentz et al., 2019). By examining youth action within norm diffusion, it is possible to acknowledge the impact of bottom-up youth initiatives in diffusing specific norms of reconciliation, namely unity and inclusivity across difference, social justice and human rights preservation, and an orientation towards a shared future. These norms are advocated by youth actors across the globe through the use of innovative tools, such as art (Dragija, 2020; Kosic & Tauber, 2010; Kurze, 2016; Peace Direct, 2019; Wollentz et al., 2019), sustained dialogue (Life & Peace Institute, 2017), and peace education (Kasherwa, 2019; Peace Direct, 2019). Their impact is visible at different stages of the norm diffusion cycle: at the national level, governments are adopting youth-related policies and including youth in representative bodies (Mengistu, 2017), as well as implementing peace education programs in educational systems (UNESCO, 2013); at the transnational level, youth transnational advocacy networks (TANs) and youth NGOs are becoming prominent and influential in promoting norms engineered by youth actors; at the international level, influence of the youth is observable in the adoption of the UNSCR 2550 on Youth, Peace and Security (YPS) of 2015, which has contributed to producing a normative shift on a global scale, enabling extensive youth participation and increasing young people's effect in diffusing their norms.

Within the norm diffusion framework, little has been done to trace the role of young people involved in reconciliation. Therefore, this research aims to illustrate the ways in which youth actors contribute to diffuse norms for the promotion of reconciliation. In other words, it seeks to delineate young people's norm entrepreneurship. In doing so, it examines the post-genocide context of Rwanda, and draws its data from 10 semi-structured interviews conducted with young Rwandans involved in the reconciliation process. The study specifically focuses on exploring the norms they promote and the strategies they use , and on evaluating their impact within and beyond national borders. The first part of the paper is dedicated to the analysis of existing norms of reconciliation in the country, as well as the methods employed by the state to promote such norms. In conformity with international norms aimed at reconciliation, Rwanda has advanced justice, truth, forgiveness, and collective memory, and additionally, unity and a common identity, through the implementation of transitional justice institutions, namely the ICTR (Des Forges & Longman, 2004; Tiemessen, 2004), the Gacaca courts (Tiemessen, 2004; Meyerstein, 2007; Rettig, 2008), the National Unity and Reconciliation Commission (NURC) (Clark, 2010; Mgbako, 2005), and the National Commission for the Fight against Genocide (CNLG) (Korman, 2014), along with the use of the education system (Russell, 2015; 2018; 2019). The second part relies upon data obtained through qualitative interviewing, and serves to identify the work of Rwandan youth as norm entrepreneurs within the

present context. First, it looks into the preconditions that determine their influential action: being the majority, having the resources, being innocents and victims at the same time, and having different, more open mindsets and perspectives. Second, it dives into the norms they promote. Here, it emerges that youth-led norms mainly pertain to two overlapping categories: the analysis of the past and the focus on the future. The former includes norms of acknowledgement (of the past) and acceptance (of the truth), while the latter involves norms of unity, a common identity, and a "never again" narrative. In between these two groups, young Rwandans also emphasize critical thinking, empathy, and responsibility. To different extents, such norms are a result of norm localization by the hands of the Rwandan new generation. Third, the research offers an overview of the methods used by youth to advance such norms. These approaches, which are clearly connected to global patterns of youth-initiated methods, are categorized into three groups: dialogue, conversation, and debate; art, writing, and storytelling; and media and technology. Fourth, the last section seeks to illustrate the concrete impact of Rwandan youth as norm entrepreneurs at national, transnational, and international levels: between norm emergence and cascade, youth-led initiatives are widely recognized within Rwandan society, especially through their extensive interaction with the state's institutions aimed at norm promotion; between cascade and internalization, youth actors are expanding their communities and forming webs of norm entrepreneurs across the globe.

This research contributes to the scholarship on norm diffusion and reconciliation. On the one hand, it sheds light on the under-recognized agency of youth actors in norm diffusion, as it permits conceiving them as norm entrepreneurs; on the other hand, it enlarges existing knowledge surrounding the role of young people in contexts of reconciliation. In terms of social relevance, this paper emphasizes the youth's ability to (re)construct social discourses in post-conflict societies, and to go beyond national borders. Consequently, it seeks to encourage the inclusion and recognition of youth agency at a national level, where young people should be equipped with the physical spaces and means to be able to achieve meaningful outcomes. Given the rising acknowledgement of youth action and potential in Rwanda, the Rwandan case serves as an example for the world.

2 Research Context

2.1 Norm Diffusion

Norm diffusion in international relations has been extensively explored on the basis of a constructivist theoretical framework offered by Finnemore and Sikkink in 1998. According to these scholars, norms, which they define as "standard[s] of appropriate behavior for actors with a given identity" (p. 891), undergo a three-stage process characterized by "norm emergence," "norm cascade," and "norm internalization" (p. 895). In the first phase, norms are built by norm entrepreneurs, who persuade their community to adopt new principles of appropriateness, with the help of organizational platforms (e.g.,NGOs); in the second phase, other countries or regions are pressured to adopt these new legitimate norms, notably through the work of international organizations and networks; in the final phase, norms are accepted and internalized by the international community, thus achieving a "taken-for-granted" character (Finnemore & Sikkink, 1998).

Of particular significance is the notion of norm entrepreneurs. Such individuals are often described as "meaning managers" or "meaning architects" (Finnemore & Sikkink, 1998, p. 897), because of their ability to reshape social meaning within broader society. In the latter, norm entrepreneurs usually encounter initial contestation, considering their use of non-normative approaches to diffuse norms (Finnemore & Sikkink, 1998, p. 897). Nonetheless, such inappropriateness is employed to ensure that their voices are heard, with the greater aim to divulge their ideas among a wider public. Finally, the motivations behind the action of norm entrepreneurs range from altruism and empathy to ideational commitment (Finnemore and Sikkink, 1998, p. 898).

Other scholars have problematized the norm life cycle by analyzing norm contestation (Stimmer, 2019; Wiener, 2004; 2017). In particular, Wiener (2004) focuses on instances of "contested compliance," while Stimmer (2019) complicates the binary understanding of contestation and acceptance by offering four distinct contestation outcomes, namely norm clarification, norm recognition, norm neglect and norm impasse.

Another substantial body of research examines norm localization (Acharya, 2004; Capie, 2008; Groß, 2015; Tholens & Groß, 2015), defined by Acharya (2004) as a process through which "norm-takers build congruence between transnational norms [. . .] and local beliefs and practices" (p. 241). Most significantly, Acharya (2004) expands on existing literature around norm diffusion by emphasizing the agency of local actors in reinterpreting and adapting outside norms to their social contexts. Groß (2015), who analyzes the construction of local meaning regarding international norms of democracy and minority rights in Kosovo, provides an example of such local agency. Finally, the role of local actors is taken forward by Acharya (2011), who develops the concept of "norm subsidiarity" in Third World countries. Here, local actors create norms with the aim of maintaining local autonomy.

Lastly, a wide range of literature concentrates on contextualizing norm diffusion, contestation, and localization, as well as exploring the role of specific norm entrepreneurs and norm-takers (Acharya, 2013; Björkdahl, 2006; Boesenecker & Vinjamuri, 2011; De Almagro, 2018; Vuković, 2020; Williams, 2009). On the one hand, Björkdahl (2006) and Vuković (2020) focus on external norm entrepreneurs, re-

spectively the UN and EU mediators. On the other hand, Acharya (2013), De Almagro (2018), Williams (2009), and Boesenecker and Vinjamuri (2011) examine internal dynamics of norm diffusion and the agency of norm-takers in shaping and localizing external norms. For instance, Williams (2009) and Acharya (2013) investigate the circulation of the "Responsibility to Protect"-norm in Africa.

2.2 Reconciliation

Reconciliation, which Bar-Tal and Bennink (2004) claim is "the necessary condition for stable and lasting peace" (p. 17), has been largely conceived as "an agreed upon norm of conflict resolution" (Rosoux, 2015, p. 48). In other words, reconciliation is internationally viewed as an essential requirement for the successful settlement of violent conflicts and the achievement of peace and stability.

Reconciliation is broadly defined by Rosoux (2017) on the basis of three levels: structural, psycho-social, and spiritual. The first indicates the development of political and economic collaboration between two parties; the second relates to the improvement of emotional understandings between the two; the last one refers to the rehabilitation of ruptured spiritual connections between victims and perpetrators.

Another significant aspect of an effective reconciliation process consists in the simultaneous mobilization of top-down and bottom-up approaches (Bar-Tal & Bennink, 2004; Rosoux, 2009; 2015; 2017) As Rosoux describes (2009; 2015; 2017), reconciliation must involve political leaders at a macro level, local institutions and NGOs at a meso level, and individuals at a micro level. Only through this threefold, comprehensive involvement of society can an effective reconciliation process take place.

A substantial amount of research focuses on the social, psychological, and spiritual level of reconciliation through the identification and problematization of the following key components: justice (Bar-Tal & Bennink, 2004; Rosoux, 2009; Staub, 2013; Weinstein & Stover, 2004), truth and collective memory (Bar-Tal & Bennink, 2004; Kelman, 2004; Rosoux, 2009; Staub, 2013), and forgiveness (Auerbach, 2004; Bar-Tal & Bennink, 2004; Rosoux, 2009; 2015; Rosoux & Brudholm, 2009).

2.3 Merging Norm Diffusion and Reconciliation

The link between components of reconciliation and norm diffusion can be perceived especially in the international normative significance attributed to justice, truth, forgiveness, and collective memory. These elements, which are generally included in the process of transitional justice, are considered as the appropriate standards of reconciliation.

Transitional justice has come to be internationally viewed as a norm of reconciliation (Ben-Josef Hirsch, 2014; Boesenecker & Vinjamuri, 2011; Kostovicova & Bicquelet, 2017; MacKenzie & Sesay, 2012; Ottendörfer, 2013; Sikkink & Kim, 2013; Subotic, 2015). Post-conflict states are expected to implement justice through the institution of different mechanisms which ensure legal accountability, with regard to individuals engaging in substantial human rights violations (Sikkink & Kim, 2013). This "individual criminal accountability model" (Sikkink & Kim, 2013, p. 276) first emerged with the Nuremberg and Tokyo trials, due to an ideological change stemming from the human rights movement (Sikkink & Kim, 2013). Indeed, these post-World War II trials laid the foundations for the so-called "justice cascade" (Ben-Josef Hirsch, 2014; Sikkink & Kim, 2013), which involved the establishment of international tribunals, such as the ICTY, the ICTR, and ultimately the ICC, for a global and decentralized prosecution of perpetrators. The justice cascade then influenced national policies across the globe, particularly through the work of TANs and peer pressure (Sikkink & Kim, 2013).

Truth is also viewed as an inherent, indispensable component of the path to reconciliation (Rosoux, 2009, p. 555). The normative and systematic relationship between reconciliation and truth can be observed in the institutionalization of truth and reconciliation commissions (TRCs), which act as transitional justice tools to expose the truth and shame perpetrators. TRCs originated in Latin America in the 1980s, and later acquired international recognition as a means to generate effective reconciliation and social reconstruction (Ben-Josef Hirsch, 2014).

Other than diffusing the norm of truth, TRCs also encourage forgiveness by hands of 'the victim' towards 'the perpetrator' (Rosoux & Brudholm, 2009, p. 34). The emergence of forgiveness as a norm is underlined by Auerbach (2004), who associates it with the influence of Christianity and Western cultures (p. 153). Today, forgiveness is advocated by governments, NGOs, religious leaders, and mediators (Rosoux & Brudholm, 2009, p. 34), and the forgiver has acquired a privileged image, considered a model of humanity and morality. (Rosoux & Brudholm, 2009).

Finally, implementing collective memory is also regarded as a crucial requirement of a successful reconciliation process. Lemarchand (2006) hints at the understanding of memory as a norm by mentioning the universal popularity of the phrase "the duty to remember" (p. 21). The diffusion of collective memory as a norm of reconciliation is especially visible in the institutionalization of museums and memory sites in post-conflict states (Sikkink & Kim, 2013, p. 270).

Nonetheless, scholars have shown how these norms, included in the transitional justice process, have encountered contestation, reinterpretation, and localization (Boesenecker & Vinjamuri, 2011; Kostovicova & Bicquelet, 2017; MacKenzie & Sesay, 2012; Ottendörfer, 2013; Subotic, 2015). Con-

testation refers to the resistance and tension created by the application of a norm in a given state (Stimmer, 2019); reinterpretation, which occurs as a consequence of contestation, implies the association of a new norm with a pre-existing norm (Acharya, 2004); localization, which also settles norm contestation practices, involves an adaptation process between international norms and local understandings (Acharya, 2004). In particular, contestation, resistance, and normative divergence have occurred in the Balkans (Kostovicova & Biquelet, 2017; Subotic, 2015), in Sierra Leone (MacKenzie & Sesay, 2012), and in Timor Leste (Ottendörfer, 2013), where agreed-upon international norms and tools of reconciliation were inconsistent with their mindsets (Ottendörfer, 2013, p. 25).

2.4 Youth Actors

In post-conflict areas undergoing processes of reconciliation and peacebuilding, youth actors are often "othered," labeled as "dangerous subjects" (Kasherwa, 2019, p. 2), "perpetrators of violence," and "problematic" individuals (Life & Peace Institute, 2017, p. 4), and youth action is dismissed, seen as irrelevant. For this reason, the youth faces multiple challenges in exercising its agency within global and national political discourses (De Graaf, 2014; McEvoy-Levy, 2006; Mengistu, 2017). Nonetheless, youth actors are gaining momentum in the context of international norms of reconciliation, as they are increasingly acting as norm entrepreneurs. Scholars have widely explored the significant role of young people in post-conflict areas (De Graaf, 2014; Dragija, 2020; Kasherwa, 2019; Kosic & Tauber, 2010; Kurze, 2016; McEvoy-Levy, 2006; Nguyen-Marshall, 2015; Prisca et al., 2012; Shipler, 2008; Thapa, 2009; Wienand, 2013; Wollentz et al., 2019). In particular, McEvoy-Levy (2006) highlights youth actors' potential by describing them as "peace resources" (p. 12) and stresses their effectiveness in promoting healing and reconciliation by virtue of their desire "to do good, to make contributions, to change systems, and to redress wrongs" (p. 21). Similarly, De Graaf (2014) underlines the importance of giving a voice and providing agency to youth actors in order to build peace and transform entire societies. The following sections serve to illustrate the concrete role of young women and men within the norm diffusion cycle elaborated by Finnemore and Sikkink (1998), in relation to reconciliation practices (see Figure 1).

2.4.1 Bottom-Up Youth Initiatives

At a local level, youth actors are increasingly contributing to the promotion of new norms of reconciliation by spreading innovative ideas and values. Early signs of youth involvement aimed at reconciliation, as well as their capacity to positively influence post-conflict discourses, were apparent during the 1960s and the 1970s, in the contexts of the German-Israeli

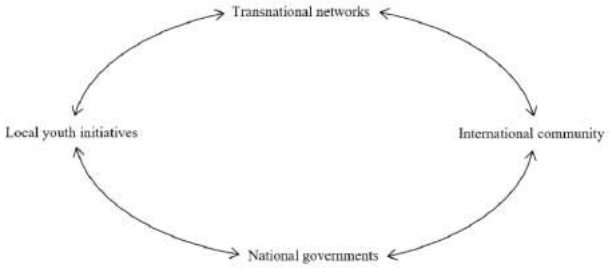

Figure 1: The role of youth within the norm diffusion cycle.

relations (Wienand, 2013) and the Vietnam War (Nguyen-Marshall, 2015). In the former, German university students initiated German-Israeli study groups; they intended to overcome prejudices, modify the conception of the "other" and envision "a joint future by means of reconciliation" (Wienand, 2013, p. 207). In the second case, students in Vietnam formed associations and organized peaceful protests to achieve social justice, fight for religious freedom, and uphold freedom of speech (Nguyen-Marshall, 2015, p. 53). In both contexts, diverse youth actors assembled and cooperated despite their differences in nationality, political affiliation, or religion, for the promotion of core ideas and values, namely unity and inclusivity, justice and human rights preservation, and an orientation towards a shared future. These three principles, in particular, acquired the character of norms as they began to be prioritized and advanced by youth actors in different contexts of reconciliation, through the use of hands-on and ingenious methods. The reasons behind a common focus on these three norms in particular are threefold. Firstly, unity across difference is emphasized as young people in post-conflict societies are often not involved in past atrocities, unlike older generations, and thus feel equipped and qualified to accept and include the "other" (Wienand, 2013, p. 206-207). Secondly, the commitment for social justice, which appears to be closely tied with human rights protection, is viewed as a priority as youth activists "are the product of a global spillover effect of international human rights practices" (Kurze, 2016, p. 2). Finally, a common commitment towards a shared future is advocated as young people see themselves as builders of a new world (UNOY, n.d.), and, by learning from the older generation's mistakes, are more prone to strive for a better future.

In the Balkans, the youth pioneered the use of artistic performances and installations as an alternative to the traditional transitional justice mechanisms (Kurze, 2016). These bottom-up initiatives, initially implemented in Bosnia-Herzegovina, were aimed at creating spaces for victim confrontation, denouncing human rights violations, and commemorating the past (Kurze, 2016). In doing so, young people in the Balkans have acted as "insider proponents" (Acharya, 2004, p. 248) of an alternative norm to the existing transitional justice norm and its implications. They attempted to "localize a normative order" (Acharya, 2004, p. 249), namely transitional justice and its tools (e.g., TRCs), by using artistic means as the alternative. This practice was

consequently amplified to the region: for instance, Wollentz et al. (2019) illustrate how young people in the divided city of Mostar have creatively rebuilt a monument dedicated to workers' rights, in which everyone could participate, thus using art to minimize ethnic divides, create a sense of togetherness, and conceive a shared future. Similarly, NGOs in Vukovar, Croatia have contributed to legitimize this practice by bringing young Serbs and Croats together through creative and artistic activities like painting and dancing (Kosic & Tauber, 2010, p. 87). The use of art within reconciliation processes soon spread to other post-conflict contexts, especially through online platforms and social media, and Balkan youth became "a role model for other youth movements across the globe" (Dragija, 2020, p. 65) For example, the Colombian youth-led organization BogotArt promotes unity and acceptance between ex-combatants and civilians in the aftermath of the 2016 peace agreement (Peace Direct, 2019, p. 22). In all these cases, youth actors prioritized the advancement of the same core values, namely unity, justice, and the visualization of a shared future, through the use of art as an original reconciliation medium.

Another innovative tool of reconciliation used by youth actors to encourage, above all, unity and inclusivity across ethnocultural barriers, is the use of peer-to-peer dialogues and interactions. An example would be the Ethiopian "Sustained Dialogue" (SD) initiative, a method of post-conflict social transformation which enabled participants to "systematically probe and gradually deal with the causes of deep-rooted human conflict" (Saunders, 2011, p. 1). The SD practice was initially applied to the Arab-Israeli-Palestinian peace process of the 1970s (Saunders, 2011), and was later extended to Ethiopian youth in 2009. Once extended to other states of the Horn of Africa, namely Sudan and Kenya, it became a tool for reconciliation among youths of the region (Life & Peace Institute, 2017, p. 10). Here, young women and men started to engage in SDs to build and improve social relationships beyond identity markers of ethnicity, religion, gender, or economic status (Life & Peace Institute, 2017, p. 8). In doing so, Ethiopian, Sudanese and Kenyan youths were able to develop mutual trust and learn to cooperate in unity despite diverse ethnic boundaries (Life & Peace Institute, 2017, p. 10-11).

Additionally, young people usher the visualization of a shared future of peace and cooperation through peace education activities. The purpose of peace education is to "promote understanding, respect, and tolerance toward yesterday's enemies" (Salomon, 2002, p. 4), as well as advancing human rights, spreading a culture of peace, and providing skills for managing and preventing violence and conflict in the future (Salomon, 2002, p. 5). Peace education programs were created and implemented in multiple contexts of intractable conflict and/or interethnic tension, but also during periods of experienced tranquility (Salomon, 2002, p. 5-6). Inspired by these programs, young people in post-conflict areas have taken it upon themselves to organize workshops and activities of peace education. This is occurring extensively across African regions: in Burundi and DRC, a myriad of youth organizations are promulgating a culture of long-lasting peace and overcoming ethnic and political divides (Kasherwa, 2019, p. 19), while in Madagascar, the youth peacebuilding organization "Act in Solidarity" has launched the program "Youth Students for Peace," which provides peace training and mentoring to university students (Peace Direct, 2019, p. 33). Through these peace education-related activities, youths across Africa and the rest of the world are divulging the importance of envisioning a shared peaceful future, marked by unity across difference and human rights protection.

2.4.2 Top-down Approaches: National, Transnational and International Levels

At the national level, governments are increasingly pressured to recognize the role of young actors in peacebuilding and reconciliation, and to involve them in social and political decision-making. In several African countries, these efforts are evident in the adoption of national youth policies and the establishment of national youth representative bodies (Mengistu, 2017, p. 3), such as the Youth and Sports Ministry (Prisca et al., 2012, p. 188). As such, youth initiatives for reconciliation are gaining ground, paving the way for new principles and methods of reconciliation to spread as national and international norms. The assimilation of such norms at a national level is particularly visible in the institutionalization of peace education in educational systems across Africa (UNESCO, 2013). Propagated simultaneously by local youths and the UN, reconciliation norms (i.e., unity and tolerance, human rights preservation, culture of peace for the future) are thus cascading to national governments, the latter increasingly persuaded to adopt these norms through the inclusion of peace education in school curricula. At the same time, the growing national attention towards youth actors enlarges young people's opportunities to reinforce the emergence and diffusion of such norms.

Youth TANs and regional NGOs are largely contributing to the circulation of norms for reconciliation developed by local youth initiatives. These organizations take the "new norms," along with innovative tools and methods, out of the national sphere and gradually push for their international recognition and incorporation into global policies as well as national governments. The United Network of Young Peacebuilders (UNOY), a TAN of 128 youth organizations operating in 71 countries (UNOY, n.d.), promotes core values like inclusivity and creativity, thus illustrating once again how norms engineered by youth actors enter norm cascade processes. In turn, this network, among many others, influences international policies.

Thanks to the extensive work of youth TANs and NGOs, the international community is pressured to recognize the po-

tential and willingness of youth actors with regard to building peace and reconciliation through the elaboration and circulation of inventive norms and tools. One of the greatest achievements at a global level is the adoption of UNSC Resolution 2550 on YPS in 2015, which states that "young people play an important and positive role in the maintenance and promotion of international peace and security" (UNSC, 2015). This resolution has contributed to increase young people's visibility in the context of reconciliation and consequently, facilitated the promotion of norms engineered by youth actors across the globe. Additionally, it has resulted in a normative shift within the international realm, as it pushes member states to enable youth participation at local, national, regional, and international levels (UNSC, 2015).

3 Methodology

The methodological approach used in this research is twofold: it consists of a case study analysis and qualitative interviewing.

The first approach is used to apply the norm diffusion framework to the context of reconciliation in post-genocide Rwanda, with the purpose of exploring which norms of reconciliation have been promoted at a national level and how the state has advanced them. In particular, the section examines the establishment of transitional justice institutions aimed at the propagation of justice, truth, forgiveness, collective memory, unity, and a common identity, as well as the use of the education system to promote a common identity, citizenship, and human rights. The analysis of these governmental mechanisms for reconciliation sheds light on the extent to which Rwanda has internalized, localized, or contested specific norms.

The second approach is employed to examine the role of youth within the norm diffusion framework of Rwanda. For this purpose, the findings of this research are based on 10 semi-structured interviews with youth actors involved in reconciliation in Rwanda. The definition of youth taken into consideration by this research derives from the National Youth Policy, according to which Rwandan youth ranges from 15 to 35 years old (Ministry of Youth, Culture and Sports, 2005). In accordance with this criterion, participants' age varies between 21 and 33 years old. Respondents, who were chosen through snowball sampling, are included in the following categories: eight of them are founders, directors, or members of youth NGOs (i.e. Peace and Love Proclaimers (PLP), Global Radiant Youth, Youth Literacy Organization, iDebate Rwanda, Iteka Youth Organization, Seven United for the Needy, and the Aegis Youth Champions program, which is part of the international NGO Aegis Trust); two of them are artists involved in artistic initiatives (i.e. Yan Events, Generation 25, African Artists for Peace Initiative); two are university students; and one is the founder of the online journal The Kigalian. All participants have a Rwandan nationality and live in Kigali, the Rwandan capital, except for two of them, who are studying abroad. The interviews took place virtually within a period of two months, and were each approximately 20 minutes long. Data analysis was conducted through thematic coding, and findings were grouped according to the following classification: preconditions for youth influence, youth norms, youth methods, and youth impact. The first part of the findings section illustrates the reasons which render them effective norm entrepreneurs of reconciliation; the second part offers an overview of the variety of norms promoted by youth actors; the third part shows the innovative ways in which young people are able to advance such norms; and the last part demonstrates the concrete impact of Rwandan youth at national, transnational, and international levels.

4 Findings/Analysis

4.1 Case Study – Norms of Reconciliation in Rwanda

In the aftermath of the 1994 genocide, where 800,000 Tutsis were massacred in just a hundred days, Rwanda embarked on a long-lasting and intricate journey of reconciliation. In line with the international norms of reconciliation, the state initiated a process of transitional justice, seeking to promote justice, truth, forgiveness, and collective memory. Additionally, the promotion of unity and a common identity was also viewed as a priority in the Rwandan context. All these norms were implemented in the country through the establishment of various institutions, namely the ICTR for retributive justice (Des Forges & Longman, 2004; Tiemessen, 2004), the Gacaca courts for truth, restorative justice, and forgiveness (Meyerstein, 2007; Rettig, 2008; Tiemessen, 2004), the NURC and the Ingando camps for unity, a common identity, and forgiveness (Clark, 2010; Mgbako, 2005), and the CNLG and memorial sites for the consolidation of an official collective memory (Korman, 2014). Another significant institution used by the Rwandan government as a tool for reconciliation is the education system, which primarily serves to propagate a common identity, citizenship, and human rights (Russell, 2015; 2018; 2019).

4.1.1 Transitional Justice Institutions

As a result of the "justice cascade" (Ben-Josef Hirsch, 2014; Sikkink & Kim, 2013), the international community pressured for the establishment of the ICTR in 1994, as a way to administer transitional retributive justice and denounce human rights violations. The ICTR, which was modeled on the basis of the already existing ICTY (Des Forges & Longman, 2004, p. 52), illustrates how the Rwandan state has internalized the process of transitional justice as an interna-

tional norm of reconciliation. Nonetheless, the tribunal faced multiple institutional challenges and turned out to be quite ineffective in its reconciliatory action (Des Forges & Longman, 2004; Tiemessen, 2004). Since it was architected by the international community, its work was not recognized by Rwandans (Des Forges & Longman, 2004, p. 62), due to widespread skepticism derived by the country's colonial past. Most notably, the main purpose of the tribunal was not in line with the need to reconcile: in compliance with the global human rights regime, the international community principally sought to "end a culture of impunity" (Tiemessen, 2004, p. 60). In turn, the focus on punishing the key criminals of the genocide overshadowed the priorities of restorative justice and healing among Rwandan society.

Another transitional justice mechanism implemented in Rwanda consisted in the local initiative of the Gacaca courts. Established by the Rwandan government in 2001 as a "uniquely Rwandan solution to a Rwandan problem" (Meyerstein, 2007, p. 503) and facing the shortcomings of the ICTR, the Gacaca courts were mainly intended to determine the truth about genocide accountability (Tiemessen, 2004, p. 62), to blend retributive and restorative justice (Meyerstein, 2007, p. 468), and to emphasize forgiveness (Rettig, 2008, p. 44). The Gacaca system was a "community event" (Rettig, 2008, p. 25), inspired by a precolonial method of dispute resolution (Karekezi et al., 2004, p. 69): a great part of Rwandan citizens were engaged as judges or witnesses in the trials. While the ICTR contributed to the internalization of the norm of transitional justice, the Gacaca courts represented a way to localize this norm to the national context, as a consequence of the ICTR failures. The Gacaca mechanism was effective as it accelerated the trial process of imprisoned perpetrators, allowing for the participation and inclusivity of Rwandan citizens, but it also presented drawbacks and criticisms. Most significantly, the Gacaca courts received international disapproval regarding its procedures, which did not comply with normative human rights laws (Meyerstein, 2007, p. 503), as well as being accused of exercising victor's justice (Rettig, 2008, p. 26). The latter refers to the wrongful prosecution of the losing party in favor of the winning party. The truth promoted through the Gacaca courts was one-sided (Des Forges & Longman, 2004, p. 63), as it was indeed based on the assumption of Hutu guilt and Tutsi innocence (Tiemessen, 2004), thus undermining impartial judgment and fostering disputes. Moreover, research has shown that a large part of Rwandans expressed concerns regarding Gacaca's competence to foster reconciliation (Zorbas, 2009; Rettig, 2008).

In addition to the aforementioned judicial mechanisms, the Rwandan government institutionalized the NURC in 1999, as a complementary tool to the justice-focused works of the ICTR and the Gacaca courts (Clark, 2010, p. 139). Through the creation of the NURC, the Rwandan state adapted the traditional model of the TRCs to its own local context, with the intention of refraining from intricate arguments about the truth (Clark, 2010, p. 141). Instead, the aim of this institution was to advance the new norm of national unity through the creation of a national identity (Clark, 2010). In order to do so, the NURC took over the Ingando initiative, which had started in 1996. The Ingando were solidarity camps in which Rwandans from all walks of life were brought together for weeks or months (Mgbako, 2005, p. 202), to learn about Rwanda's history and issues, and about their rights and duties as Rwandan citizens (Clark, 2010, p. 139). Despite these efforts, the urge to propagate national unity encountered complex problems derived by the impossibility to define oneself in ethnic terms in the aftermath of the genocide (Clark, 2010): ethnic distinctions were indeed legally obliterated by the Rwandan government after the conflict (Clark, 2010). The paradox of internalizing a common identity based on the denial of ethnicity thus resulted in a superficial unity (Clark, 2010), on the basis of a general difficulty to identify oneself. Moreover, the Ingando solidarity camps were criticized for alleged political indoctrination (Mgbako, 2005).

The state also instituted several memorial sites across the country to promote an official collective memory. These memorials were constructed according to the model of those created after the Holocaust (Korman, 2014), thus demonstrating Rwanda's internalization of the norm of memory. The normative character of memory was further authenticated through the establishment of the CNLG in 2008, which took charge of the propagation of memory across the country (Korman, 2014, p. 98). The diffusion of an official collective memory brought about certain difficulties such as the overgeneralization of the history. The possibility of remembering was only granted in relation to the experience of the Tutsi, who qualified as victims, as opposed to the Hutu, who were generalized as perpetrators, and whose memory were then forbidden by law (Lemarchand, 2006). The limitations of the Rwandan official memory are still visible during Kwibuka, the annual commemoration period, where remembrance seems to be legitimate only for Tutsi survivors (Baldwin, 2019).

Lastly, forgiveness was promoted in Rwanda as a necessary norm of reconciliation. Influenced by international NGOs and the Church, government officials, the president, and institutions like the NURC and the Gacaca courts all advocated unconditional forgiveness as an essential requirement to move on (Brudholm & Rosoux, 2009; Rettig, 2008). However, many survivors resist this norm, refusing to forgive and understand perpetrators as they lack energy to express empathy, which is viewed as an unimaginable possibility (Brudholm & Rosoux, 2009, p. 45). Other research shows that forgiveness is not viewed as an indispensable requisite for reconciliation in Rwanda (Zorbas, 2009, p. 134).

4.1.2 Education

The role of education in promoting norms of reconciliation has been widely examined by Russell (2015; 2018; 2019). The scholar examines how the Rwandan government uses the education system to advance a common identity (Russell, 2019), global citizenship (Russell, 2018) and a human rights discourse (Russell, 2015). In addition to these norms, education also acts in line with transitional justice institutions and serves to promote an official truth and an official collective memory (Russell, 2019).

Firstly, the education system is employed to "propagate a new sense of national unity" (Russell, 2019, p. 55), through the diffusion of common identity as a norm. This newly formed national identity excludes ethnicity, which is taught to be "an identity externally imposed on Rwandans, rather than as a precolonial characteristic" (Russell, 2019, p. 73). The complete eradication of ethnic identity, as well as the impossibility to discuss and explore identity in terms of ethnicity, led to the adoption of different identity indicators based on the experience of one's family during the genocide (Russell, 2019, p. 76). For instance, Rwandan youths started to define themselves as children of génocidaires or children of survivors (Russell, 2019, p. 76), thus implying further divisions on a societal level.

Secondly, education also incorporates models of global citizenship for the formation of the Rwandan identity (Russell, 2018). The international influence on the newly developed notion of citizenship in Rwanda is especially noticeable in the civic education curriculum, which involved the help of UNESCO and UNICEF (Russell, 2018, p. 390). As a result, the curriculum includes global norms related to human rights, with an emphasis on gender equality (Russell, 2015). Nonetheless, these norms are not merely internalized, but are edited and adapted at the local level (Russell, 2015). In particular, the human rights discourse is "generally mentioned in reference to gender equality rather than to ethnic groups" (Russell, 2015, p. 608). This strategic modification is contingent to the eradication of ethnic markers in the Rwandan identity.

Finally, the education system is used to promote official narratives of what represents the truth and who should be remembered through collective memory (Russell, 2019). In the teaching of history, truth and memory are molded in a way that leaves no space for discussion and debate of the past. This way, "the state has suppressed critical thinking" (Russell, 2019, p. 25) and the experience of young people coming from a perpetrator background is fully disregarded (Russell, 2019, p. 135). These individuals are consequently denied the chance to reflect on their own experience of pain.

4.2 Interviews – The Role of Youth

In Rwanda, the youth is increasingly affecting the discourse surrounding norms of reconciliation. The following sections, which are based on data collected through 10 semi-structured interviews with youth actors, illustrate the role of youth within the existing framework of norm diffusion in the country. The first part explores the preconditions that allow young people to exert influence, and that inform the kinds of norms that they promote; the second dives into the multitude of norms of reconciliation diffused by Rwandan youth, and shows how these norms are interlaced; the third focuses on the innovative methods and approaches employed by young people to spread such norms; and the last demonstrates the tangible impact of the youth's actions at the national, transnational, and international levels.

4.2.1 Preconditions for Youth Influence

Before examining the variety of norms of reconciliation promoted by Rwandan youth, it is necessary to explore some of the factors which determine their ability to exert influence. It emerged from the interviews that young people's power to diffuse norms is induced by the following elements: first, they represent the majority of the population in Rwanda; second, they have more and better resources than the previous generation; third, they are all innocent, by virtue of their non-involvement in the genocide, but at the same time they are all victims of their parents' heartaches; finally, they offer a different perspective and an open mindset, once again due to their lack of participation in the 1994 atrocities.

When asked about the significance of young people in reconciliation, most participants mentioned the youth's percentage of the population in the country. Rwandans below the age of 25 make up 67% of the total population (Ministry of Youth, Culture and Sports, 2005), reaching almost 80% if considering Rwandans between the age of 25 and 35. This aspect enhances young people's potential to have an impact on society, while also giving the incentives, at a national level, to include them in the political sphere and to support their programs. In Rwanda, there is a strong awareness related to the plurality of young people and their intrinsic importance for the country:

> "Youth in Rwanda are the majority of the country. So, leaving them behind would actually mean leaving the country behind."
> – Shema, PLP executive director

Another precondition is the abundance of resources. In terms of material resources, young people are more exposed to the Internet, which grants them the possibility to carry out more research. They have the ability to gather information, but also to discuss it with their peers in a language

that they can all understand. These resources indeed enable young people to better connect to each other and engage in dialogue, thus fostering the diffusion of ideas:

> "For example, there are kids in their house that are told that all of these things [genocide events] are a lie, from the beginning. [...] When young people are saying these things, they are telling them to their peers, and peers have the ability to question. They're like, okay, this is not what I heard, this is what's been happening. And then you know, it creates those spaces. [...] And they can actually speak a language more young people can understand. And for things like reconciliation to work, it needs to be understood."
> – Yannick, professional dancer

In addition, the youth has the energy to initiate, explore, debate, and understand. This resource places them in a favorable position to develop empathy and engage in conversation. On the contrary, the previous generation of survivors is conditioned by an "immense fatigue" (Brudholm & Rosoux, 2009, p. 45), which prevents them from listening to and comprehending the 'other'.

Since they were not involved in the genocide, young people ascribe to themselves a condition of innocence, which in turn, determines the fact that they carry less emotional weight than their parents' generation:

> "The new generation are innocent. Because they never participated in the in the politics or ideology which led to the genocide."
> – Dieudonné, Global Radiant Youth founder

> "We are in a unique position where we know the wounds and the pain that history has caused, but we don't feel it as deeply as our parents do, which means that we have a chance to also do the work that [...] allows us to create new knowledge for other people, in a way in which some of these ideas affect people."
> – Jean Michel, iDebate and PLP founder

Nonetheless, they also consider themselves as victims of a legacy they did not create. This shared circumstance allows them to acknowledge that everyone, despite their identification as children of survivors or perpetrators, carries on their parents' traumas in different ways. In doing so, they are able to overcome the binary division between the survivor identity and the génocidaire identity which usually evolves in schools (Russell, 2019, p. 76), and which undermines unity.

Finally, their detachment from the wounds of the past allows them to offer a different perspective and a generally more open mindset. On the one hand, the fact that young people have mostly experienced a peaceful environment gives them the advantageous possibility to provide fresh ideas, as well as more positive and liberal perspectives. On the other hand, the distance between them and the genocide events leads them to acquire an unbiased, open mindset which is free of stereotypes and "us vs them" narratives. In turn, this allows them to analyze the past and question assumptions in a critical and objective way:

> "We have an opportunity, we have a negative past and now, we have to compare, and then we'll have the opportunity of making a choice. I think there's a previous generation who never had this opportunity because they could only see a single side, a single story of thinking of reality, of the history, a single way to look at things."
> – Dieudonné, Global Radiant Youth founder

> "When you didn't go through it live, there's a lot of things you can question. [...] And younger people are more open to that, younger people are more open to be like, okay, what's happening? What is this? How did it start? Why did you guys do this?"
> – Yannick, professional dancer

4.2.2 Youth Norms

The norms of reconciliation diffused by Rwandan youth actors can be divided into two overarching and overlapping categories: the analysis of the past and the focus on the future. In relation to the former, young people promote acknowledgement (of the past) and acceptance (of the truth). Included in the latter are the norms of unity, a common identity, and the "never again" narrative. Furthermore, they emphasize critical thinking, empathy, and responsibility, which serve both to examine the past and to shape the future. Figure 2 illustrates the interlaced character of youth norms.

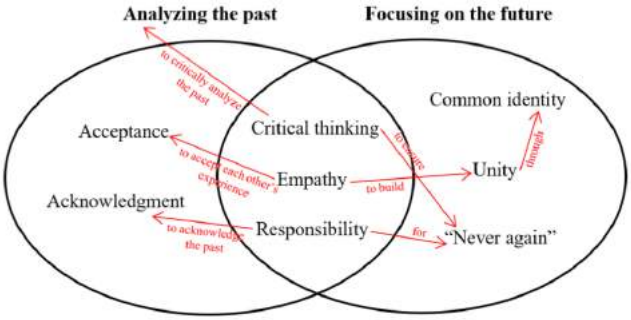

Figure 2: Youth norms

Analyzing the Past

> "Priority is on the root of things, not on the leaf of symptoms."
> – Shema, PLP executive director

As illustrated by the quote above, part of the work of the Rwandan youth as norm entrepreneurs consists of analyzing the past. Young people propagate the necessity of learning from past mistakes as an essential requirement to achieve reconciliation:

> "When we train people, what we do is we go back into history, and look where things did not go well."
> – Marc, PLP founder and Aegis Youth Champions Program manager

The value attributed to the past is necessarily influenced by norms of truth and memory promoted by the state through the education system (Russell, 2019). In schools, students are taught about the 'official' history of Rwanda and the truth about the genocide, and are given an 'official' collective memory to uphold, but are not offered the chance to discuss and critically examine these events (Russell, 2019, p. 25). Rwandan youths today are actively exploring and questioning past actions, as they utilize critical thinking to understand the past and to avoid repeating the same mistakes:

> "When you look at one of the things that most perpetrators say, is that the reason why they did something is because they were taught by people in leadership to do it. So, we've had a culture that is very obedient. [...] We are creating a culture in a way that is the opposite of the kind of culture that was there before. [...] I'm hoping that once we have young people who have the critical thinking skills, [...] they will be able to think for themselves."
> – Jean Michel, iDebate and PLP founder

Nonetheless, the youth is careful in ensuring the acceptance of the truth:

> "I think that when you accept what happened, you see the bigger view in front of you, [which allows you] to move on with your life."
> – Sandra, university student

While the norm of truth encounters contestation among the previous generation of survivors and perpetrators, as it is often perceived as single-sided (Des Forges & Longman, 2004; Tiemessen, 2004). However, the youth is more careful in framing it in such a way that involves the recognition of the experience of the "other." Thus, on the one hand, they strive for the acceptance of historical facts, in line with the narrative fed by the government through institutions; on the other hand, their acceptance does not only accommodate young people with a survivor background, but also addresses the hardship of youth coming from a perpetrator background:

> "We are working with children from all families, so it's regardless of who is who."
> – Samuel, Seven United for the Needy and Iteka Youth Organization member

In order to do so, young people promote empathy as an essential skill:

> "Understanding people's experience, even though history is there, but people experienced it differently, right? And it hurts each person differently, meaning we don't see one hurt being bigger than the other. [...] So, it takes humility to accept that history happened to you differently. And then, because history happened to you differently, it allows me to also listen to you."
> – Jean Michel, iDebate and PLP founder

In relation to memory, young people question their responsibility to remember something they did not experience, and instead, they diffuse the responsibility to acknowledge the past, in order to tackle genocide ideology in the country and abroad. This way, the contested "duty to remember" (Lemarchand, 2006), which only applies to the experience of Tutsi survivors, evolves into the responsibility to acknowledge all kinds of experiences, thus overcoming the limits of the existing official collective memory.

> "The idea of remembering is also a very key thing that we talk about: how do you remember something you've never experienced? [...] There's a lot of, you know, alternative stories, they say it was not a genocide, it was a double genocide, it was war, it was this, it was that. And the idea of continuously [...] looking back at something you haven't experienced, because most of us didn't necessarily see the genocide, but remembering, putting your voice to it, like saying, okay, this is what happened, [...] this is my responsibility towards the country today, the society today, myself today, my family today. And, to be honest, that is it, that is commemoration, that is remembering, to us."
> – Yannick, professional dancer

Focusing on the Future

> "We are a bridge between [the older] generation and our children. And if we do not take this responsibility now, that means you're going to carry on [the violence] in your hearts, and then you're going to transfer it to our children, which means the circle will never end. So, we need to end the circle of violence now. [...] I take the bitter pill now as a young person, [...] so that my children can live better."
> – Marc, PLP founder and Aegis Youth Champions Program manager

> "There is a power I have, to shape the future."
> – Dieudonné, Global Radiant Youth founder

Youth responsibility does not only involve acknowledging the past, but also, and most importantly, committing to a brighter future. All respondents stressed their duty to build a better society for them and for future generations, a society that is free of violence and conflict. In order to do so, they promote a "never again" norm. Here, youth actors seem to have a more powerful word on the subject, as they prioritize their responsibility to shape the future rather than remembering the past, which they acknowledge and analyze with a critical mind, to avoid history repeating itself.

> "Our impact is stronger than our parents' on the "never again" story."
> – Sandra, university student

The prevailing norm advocated by the youth in view of the future is unity:

> "I'd say that 70% of the youth in Rwanda is aware and is on the fight for [...] full unity. [Young people] are teaching unity, they're pressing for unity and reconciliation, they're into it."
> – Eunice, university student and PLP committee member

This emphasis on unity is clearly induced by the state's efforts to promote national unification through the propagation of a common identity (Clark, 2010). In particular, young people are taught in school about their pre-colonial, non-ethnic, shared identity (Russell, 2019), and they uphold and diffuse it as a complementary element to the notion of unity. As shown by the following quote, their understanding of ethnicity and identity is in line with what they learn in school, namely that ethnic markers were only social constructions imposed on them by the colonizers (Russell, 2019, p. 73):

> "If you look back into the history that led to the genocide, we used to have what became ethnicity, was once social classes, you know, it wasn't ethnicity, but then the colonizers and the missionaries, because they wanted to divide and rule, converted what was social classes into ethnicity and then started dividing people."
> – Shema, PLP executive director

As a consequence, they push for the recognition and internalization of this pre-colonial common identity, as a way to unify Rwandan society. While endorsement of this norm might have been painful and complicated for a previous generation affected by the "us vs them" narrative (Clark, 2010), today's young people, influenced by a condition in which ethnicity is meaningless, are more prone to explore the meaning of this newly-developed identity:

> "What is our identity? What do we become when you're born after something like that? How do you identify yourself? From your parents being classified Hutu, Tutsi or Twa or whatnot, then seeing where it led them. Then also understanding that it was all something that they learned, something that they were taught. Now, how do you come back and be Rwandan today? What does it mean? Like, what does it feel like to be Rwandan today? What are the responsibilities? What goes with this new identity? Well, it is not new, because we've always been Rwandan. But this is a new way of looking at Rwandanness."
> – Yannick, professional dancer

In consonance with global norms of citizenship and human rights that students learn in school (Russell, 2018), the "new Rwandanness" seems to imply a strong openness and acceptance towards differences. As demonstrated by the following quotes, young Rwandans today consider themselves first as human beings,and view their contrasting backgrounds or characteristics as a positive asset to unity:

> "To promote the value of humanity means to treat all human beings with respect, fairness, and dignity, independently of their age, gender, nationality, religion, or background. [...] If you have that, the value of humanity, whatever happened to you, you still have this courage to know that we are still human beings. Yes, we can, [...] we need to remember that we are all humans."
> – Yannick, Iteka Youth Organization and Yan Events founder, professional artist

This way, youth actors are able to move beyond the "fragile, superficial unity" (Clark, 2010, p. 144) propagated by the government and the state's institutions. Their notion of unity is reframed, as it is not only based on identity, but informed

by the idea of humanity and by their empathic skills. These elements allow young people to disseminate a more sustainable and resistant unity for a future society.

> "[Reconciliation happens when] they start developing empathy for one another, when they start looking at each other as a human being more than what they went through in their previous experiences. And then eventually, that builds a bigger community [...] that people can share."
> – Marc, PLP founder and Aegis Youth Champions Program Manager

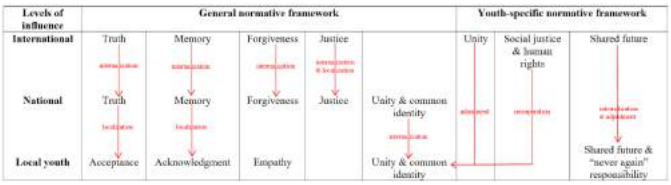

Figure 3: Interplay between preexisting norms and youth norms

This section has illustrated that, overall, most of the work of Rwandan young people as norm entrepreneurs consists in the localization of existing international and national norms (see Figure 3). Most notably, "the trajectory of localization" (Acharya, 2004, p. 251) applies to the existing norms of truth and memory, which are redefined respectively as acceptance and acknowledgment. In the "prelocalization" phase (Acharya, 2004, p. 251), local youth actors resist and contest the limited, single-sided normative truth, as it does not account for the experience of those coming from backgrounds other than the 'survivor' one. Similarly, they do not fully endorse the norm of memory, as the notion of remembrance is incompatible with a generation that was not directly involved in the genocide events. In the second phase, local actors "frame external norms in ways that establish their value to the local audience" (Acharya, 2004, p. 251). In relation to truth, young Rwandans expand its applicability to the whole of society, through empathy and critical thinking. Similarly, memory is reframed to accommodate the experience of the post-genocide generation. In the "adaptation" phase, youth actors "redefine the external norm" (Acharya, 2004, p. 251): truth is reformulated as acceptance, while memory is redefined as acknowledgment. Moreover, the widely contested norm of forgiveness is not necessarily addressed by the Rwandan youth, who concentrate on spreading empathy instead.

Other norms are internalized with slight modifications. In line with global youth trends regarding an orientation towards a shared future, Rwandan youth internalize the responsibility for a nonviolent future, with a stronger focus on the "never again" norm. Furthermore, the norms of unity and common identity, which are heavily encouraged by the Rwandan government, are not necessarily localized by youth actors, but seem to be propagated "to enhance the legitimacy and authority of their extant institutions and practices" (Acharya, 2004, p. 248). In agreement with the education system and state's institutions, the new generation fully upholds the Rwandan identity and the need for unity. Their views are informed both by global citizenship ideals present in Rwandan school curricula and international youth norms of social justice and human rights protection. These young Rwandans' understanding of unity does not only entail the notion of "Rwandanness," but a humanistic openness towards differences beyond national boundaries. Thus, in essence, Rwandan youth is largely contributing to shaping global norms of reconciliation into "international-local hybrids" (Tholens & Groß, 2015, p. 251) on the ground.

4.2.3 Youth Methods

When asked about the different kinds of methods they use to foster reconciliation, interviewees revealed a multitude of innovative and original approaches, which might be classified into three groups: dialogue, conversation, and debate; art, writing, and storytelling; and media and technology.

Firstly, these quotes show how youth actors organize dialogues, conversations, and debates, as well as create the spaces to do so:

> "We have a voice festival where people just share their ideas, their thoughts, their voices for them to be heard by decision makers and policy makers and also promote their thoughts in rural areas."
> – Dieudonné, Global Radiant Youth founder

> "We use what we call participatory methodology. And this has been really working for us, because when you engage with someone, and they also engage with you, it gives them space, it gives them a feeling that they're being heard. [...] So, we're not doing lectures, [...] we would gather around and, you know, bring speakers around as well, but also try to initiate discussions and not lectures. That way, people get to tell you what's going on in their lives and what's not going, the issues they have, they are free to talk. We're trying to create a safe space for young people to talk and have a conversation."
> – Shema, PLP executive director

By using these techniques, young people stimulate critical thinking and empathy, in order to spread norms of acknowledgement and acceptance. By coming together, discussing, and sharing ideas, the youth is invested in looking back and examining the past, in such a way that everyone gets the chance to learn the truth, to critically reflect, and to recognize and support each other's experiences:

> "I started working with my peers, with people of my age, creating groups of conversation. It has

> proven that most of us don't have this type of conversations with our parents. Our parents are literally broken, you know. [...] There's so many things happening that it is easy to think that a lot of things have been covered, and young people really used to look at it from afar. And now, my contribution personally was trying to have more of my people, people of my age, people that were born after the genocide, to really go into thinking about it, having conversations around it."
> – Yannick, professional dancer

In this context, a meaningful example is the iDebate organization, which brings debate activities and competitions to Rwandan schools and abroad, thus "encouraging a young generation of critical thinkers" (Queen's Commonwealth Trust, 2018) in view of the future.

Another original category of methods to diffuse reconciliation norms includes the use of art, which involves dance, music, theater, and painting. The Rwandan artistic scene is particularly prolific in addressing reconciliation through the advocacy of unity, empathy, and identity, other than reflections surrounding the past. One example is the Generation 25 initiative of 2019, a youth-led play which represented the experiences and burdens of the post-genocide generation and brought together diverse Rwandan youths (Ikilezi, 2019).

> "I have been trying to use art as a tool for humanity and to promote peace around the continent, in the country, and around the world. [...] Art can speak loud. And it can effectively reach as many people as possible."
> – Yannick, Iteka Youth Organization and Yan Events founder, professional artist

> "Art is the medium, art is the channel. You know, there's a lot of things that will be hard to bring up in a normal conversation, in a speech type of style. But, I mean, art touches not only the brain, but also the heart, you know, and sometimes that's what you need."
> – Yannick, professional dancer

However, the use of artistic means is not only reserved for artists themselves: it is also implemented in the programs of other youth-led organizations and initiatives. For instance, PLP offers an artistic arena, the Anda Arts Festival, for free expression and unification, while the Aegis Youth Champions Program also organizes activities regarding painting, acting, and music.

Secondly, writing is encouraged among Rwandan youth as a way to explore different ideas and, above all, to cultivate empathy:

> "Literacy and communication breed empathy. And, to me, that's central to the power of literacy, of being able to read, so the more I read, for example, what you've written, it's me trying to put myself in your shoes, looking at the world from your point of view. And literature does that. And the more I do that, the more I see the world from your perspective, it breeds empathy. So, I think it's an important step in getting people to talk to each other, feel for each other, understand each other, [...] and to reconcile."
> – Gilbert, Youth Literacy Organization and The Kigalian founder

An example of such writing initiatives is that of the Youth Literacy Organization, which seeks to spread literacy among the youth:

> "We are one of the very few pioneers of encouraging, especially young people, to write about the genocide, and write about the experiences, even if you were born after the genocide, but how did you experience it? How did your community experience it, what do you think about this whole commemoration events?"
> – Gilbert, Youth Literacy Organization and The Kigalian founder

Thirdly, young people use storytelling as a tool to promote empathy and responsibility towards the future:

> "By using these stories, we show them those steps, where people who were different, people who hated one another, take one step, a step of responsibility. And because they've taken the step of responsibility, it opens up all the other things, and eventually leads into a community. So we choose to use the storytelling approach, because people relate to stories, and they can feel the emotions in the story, they can feel the empathy in the story, they can see the critical thinking in the story."
> – Marc, PLP founder and Aegis Youth Champions Program manager

The final type of methods through which young people are able to make their voices heard is the use of digital technologies and social media. Within the online realm, they organize virtual dialogues, share their artistic performances, and write their thoughts and ideas. Especially in times of pandemic, young people are taking the lead in finding alternative, digital reconciliation activities. For instance, PLP has weekly virtual conversations, as well as online seminars and conferences, such as the Oath conference, with young people from Rwanda and abroad. In addition, The Kigalian is an independent, youth-led, future-oriented online journal in which

young writers publish articles and stories related to different topics. One final peculiar example is the Aegis Youth Champions Program's interactive voice system, which is explained in the following quote:

> "IVR is an interactive voice response system. [...] So what we did was to create those stories, put them on a phone line platform where people can dial in, and then they listen to the stories, and it can ask them questions, and then they interact with it. So that shows us which level people are at, [...] how they understand [reconciliation]. [...] Because I said, it's a process, so people will be at step 10, when others are still at step one. [...] We divide them in groups, we have the content that is for young people, but also we have those who are in school and those who are out of school, so that it can reach their context. But also, we have parents, decision makers, and we also have teachers, and all this is the same content, but delivered differently to these people so that it can reach that context."
> – Marc, PLP founder and Aegis Youth Champions Program manager

The results provided in this section highlight the link between methods of reconciliation used by Rwandan youth and global youth-initiated models to bring about peace and reconciliation (see Dragija, 2020; Kosic & Tauber, 2010; Kurze, 2016; Life & Peace Institute, 2017; Peace Direct, 2019; Wollentz et al., 2019), thus hinting at the normative character that they have acquired at a global level. In particular, the role of artistic tools and media is increasingly recognized in the country, as they are being utilized both as an alternative to official methods (i.e., transitional justice institutions) and as a complement to them (CNBC Africa, 2019). The potential of such innovative methods has been largely explored in other settings (see Dragija, 2020; Kosic & Tauber, 2010; Kurze, 2016), but is reinforced specifically in the Rwandan context, where media and art played a crucial role in the 1994 genocide. As young people dive into the analysis of the past, they learn about the ways in which perpetrators created music, paintings, and poems, and used radio stations and newspapers to spread hate and incite killings (CNBC Africa, 2019), thus understanding the significance of such tools in their society. On the basis of this knowledge, Rwandan youth is reclaiming these mechanisms for opposite purposes, namely the promotion of norms of reconciliation.

4.2.4 Youth Impact

The impact of Rwandan youth as norm entrepreneurs of reconciliation is especially visible at a national level, where they are gaining increasing recognition by the government, appearing on national television and other media, and engaging with the education system. PLP are notably influential:

> "[Our first campaign] was on radio. It was everywhere. We got support from the government and we got support from the Truth and Reconciliation Commission."
> – Marc, PLP founder and Aegis Youth Champion Program manager

> "We've been having an impact. In the commemoration period, we have been going from tv stations to tv stations."
> – Eunice, PLP committee member and university student

In particular, one of their activities, called Walk to Remember, has been instituted as a national tradition. Conceived in 2009 by the PLP founders, Walk to Remember occurs every year during the commemoration period, and it involves a large deal of Rwandan citizens, decision-makers, politicians, people from abroad, and even the president:

> "Walk to Remember is a walk which is done on the occasion of commemorating the victims of the 1994 genocide against Tutsi, and basically the lessons that we want to give to young people who attend Walk to Remember, all Rwandans who attend or even foreigners, is that when you walk you have a source and a destination, and we have a past, as Rwandans we have a past, however bad it is, it is ours, but we also have a destination, we also have a future. [...] You have a past, you cannot erase it, but you can learn from it and use that to reach your destination. [...] It has now become a national thing. And it's those things that cannot stop for any reason. [...] [The president] always comes, and we have different conversations with the president."
> – Shema, PLP executive director

Other nationally prominent youth-led initiatives are artistic performances. In 2019, the Generation 25 play was staged multiple times in the country, premiering at the Kigali Genocide Memorial (Opobo, 2019), and acquired momentum online, on social media and virtual newspapers. Finally, youth actors are growingly shaping their national context by implementing their activities in national schools: PLP members are often students who form clubs in their high schools and spread the word. iDebate, which operates within a rising number of national high schools, illustrates how norms and ideas initiated by Rwandan young people are cascading to neighboring countries: the organization has effectively implemented debate programs in schools in Uganda, Kenya, Tanzania, and Burundi, and has started an East African debate

competition. Their vision is to create an "iDebate Africa" (Queen's Commonwealth Trust, 2018). This is only one example of how Rwandan youth is reaching the international level, spreading their ideas to young people all over the world. With an open-minded, global orientation, they are diffusing unity, the "never again" narrative, and genocide acknowledgement across the globe:

> "PLP has more than 3000 members in Rwanda and around the world."
> – Eunice, PLP committee member and university student

> "We thought, what if we're able to take the story of Rwanda to the world? So, we went to the US, we spoke for almost 3-4 months, we would go doing a speaking tour and get students to debate against different people."
> – Jean Michel, iDebate and PLP founder

> "We now have communities in the US, which is now registered as friends of Seven United. We have a community in China, [...] we have friends of Seven in the UK."
> – Samuel, Seven United for the Needy and Iteka Youth Organization member

As a further example, the Oath conference was a PLP-led online event held during the 2021 commemoration period, which involved several Rwandan and international young people and organizations, as well as representatives of national institutions like the CNLG and the NURC - around 500 people in total. This virtual conversation was open to anyone who had the intention of learning and talking about the past, other than discussing young people's issues for the future. Multiple projects have been organized by Rwandan youth with the aim of spreading their voices internationally: the Walk to Remember has also become a popular initiative among youths in diverse areas of the world, representing the PLP's greatest achievement. In 2014, the event was organized in 60 cities across the globe for a total of 60,000 participants (The Independent, 2018). In relation to artistic programs to diffuse youth's norms, the African Artists for Peace initiative is an African movement invested in spreading a culture of peace and unity through creative means in the whole continent. Generation 25 also went beyond national borders as it was staged in New York and involved international artists (Opobo, 2019).

Findings related to the impact of Rwandan youth at national, transnational, and international levels demonstrate how their norms are entering the normative realm of reconciliation (see Figure 4). Between norm emergence and cascade, youth actors strategically persuade the national government and the state's institutions to adopt their new standards. Within this "highly contested normative space" (Finnemore & Sikkink, 1998, p. 897), youth entrepreneurs frame norms and methods of diffusion to render them appropriate in regard to existing standards. For instance, the implementation of debate programs in Rwandan schools was possible as a result of an adaptation process, which served to present those initiatives as if they were in the best interests of the education system and broader society. The following quote explains such adaptation process:

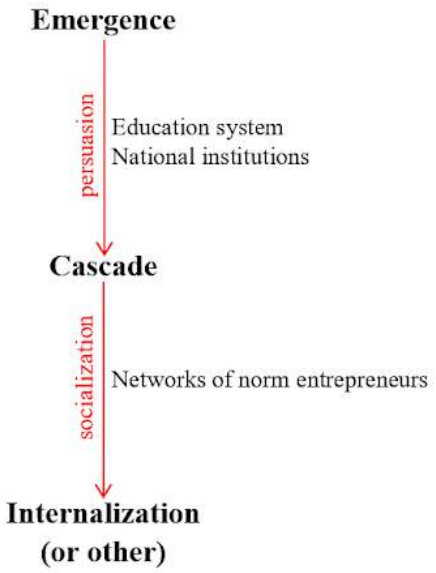

Figure 4: Youth impact within the normative realm of reconciliation

> "At first you have a lot of people who are very skeptical, and it is as if you're training young people to be rebels, or that you're training them not to listen. [But] when they saw that these young people were more articulate, that they were smarter, and that they were also doing well in university, then I was able to convince them. I call it audience adaptation. Many of them were not really interested in the critical thinking element, but they were really interested in the communication aspect."
> – Jean Michel, iDebate and PLP founder

Through their engagement with the education system, youth actors are increasingly shaping one of the government's most influential tools for diffusing norms of reconciliation. Additionally, with the purpose of "invoking a logic of appropriateness" (Finnemore & Sikkink, 1998, p. 898), Rwandan youth also acts in conjunction with official national institutions aimed at reconciliation: for example, the Walk to Remember is now annually organized with the help of the CNLG (Peace and Love Proclaimers, n.d.), while the Generation 25 premiere at the Kigali Genocide Memorial in 2019 implies that such institution, whose aim is to advocate official collective memory, has embodied youth-led artistic methods of norm diffusion. This way, youth norms are acquiring legiti-

macy at a state level.

Between norm cascade and internalization, the diffusion of Rwandan youth's norms at a transnational and international level is occurring through a process of "socialization," which is "the dominant mechanism of a norm cascade" (Finnemore & Sikkink, 1998, p. 902). Through this mechanism, young people form "networks of norm entrepreneurs" (Finnemore & Sikkink, 1998, p. 902) as they enlarge their communities to reach multiple countries, as illustrated in the above section. In turn, all these networks strengthen the validity of norms of reconciliation initiated by the new generation of Rwanda, as well as the legitimacy of their innovative methods, thus paving the way for a new normative framework of reconciliation.

5 Conclusion

This research has illustrated the role of youth actors as norm entrepreneurs within the context of reconciliation, through the analysis of the Rwandan case. In post-genocide Rwanda, the state has advanced internationally-recognized norms of reconciliation, namely justice, truth, forgiveness, and collective memory - which may all be included in the overarching norm of transitional justice -, but also prioritized new norms of unity and common identity. Such norms have been promoted through the establishment of transitional justice institutions: the ICTR exemplifies the internalization of retributive justice (Des Forges & Longman, 2004; Tiemessen, 2004); the Gacaca courts were a localized tool for restorative justice, truth, and forgiveness (Meyerstein, 2007; Rettig, 2008; Tiemessen, 2004); the NURC and the Ingando camps emerged to diffuse unity, a common identity, and forgiveness (Clark, 2010; Mgbako, 2005); and the CNLG and memorial sites demonstrate the state's internalization of an official collective memory (Korman, 2014). In addition, the Rwandan government has used the education system as a tool to spread unity, common identity, truth, and collective memory, as well as human rights and global citizenship (Russell, 2015; 2018; 2019). Despite these efforts, norms of reconciliation encountered significant challenges within society: truth and memory were perceived to be one-sided (Baldwin, 2019; Des Forges & Longman, 2004; Lemarchand, 2006; Tiemessen, 2004), the Rwandan identity was dependent on the denial of ethnicity, thus leading to the diffusion of a superficial unity (Clark, 2010), while unconditional forgiveness was seen as an impossible option for an exhausted generation of survivors (Brudholm & Rosoux, 2009).

In this context, data obtained through semi-structured interviews with young Rwandans unveiled how Rwandan youth is progressively engaging with and impacting the national norm diffusion framework aimed at reconciliation. To begin with, their demographic value, their access to resources, and their non-involvement in the genocide place them in a unique position to exert influence, as these preconditions allow them to overcome multiple challenges faced by the previous generation. Most importantly, Rwandan youth localizes existing national and international norms of reconciliation. The state's norms of truth and memory are indeed localized, thanks to the promotion of empathy, which substitutes the contested norm of forgiveness, as well as critical thinking. These norms are reframed respectively as acceptance of the truth and acknowledgment of the past, in order to incorporate everyone's experience despite survivor or perpetrator backgrounds. Redefinition of such norms allows for a wider inclusivity and a more effective norm promotion, as opposed to the work of existing institutions. In consonance with broader global youth-led norms, Rwandan youth especially propagates norms aimed at the conceptualization of a shared and peaceful future, marked by a "never again" narrative: they internalize national norms of unity and a common identity, but also incorporate notions of humanity and human rights, thus developing a stronger, more meaningful unity, which appeared to be unattainable by the older generation. Furthermore, findings illustrate diverse innovative methods used by youth actors to promote their norms: influenced by global youth initiatives, they use dialogue, debate, and conversation; art, writing, and storytelling; and media and technology. Such tools are gradually being embedded within traditional mechanisms used by the state, thus acquiring more and more legitimacy with time. Lastly, the impact of Rwandan young people within and beyond national borders highlights their involvement in the norm diffusion cycle. As they integrate their initiatives within the national education system and engage with national institutions aimed at reconciliation, they persuade Rwandan society to adopt new appropriate standards. On top of that, the Rwandan new generation is spreading their norms abroad through the creation of networks, which contribute to enhancing the validity and recognition of youth norms of reconciliation.

In light of these findings, this paper sheds light on the relevance of the youth in post-conflict areas, in which they have the potential to reconstruct and shape societies by engaging in norm diffusion. The case of Rwanda, in which young people are active participants in the reconciliation process, represents a model for other contexts torn by atrocities and divisions. Nonetheless, the present research is based on a restricted amount of data, which was collected remotely. In that regard, it acts as a starting point for broader, on-the-ground analyses of youth actors involved in reconciliation, and on the whole in the framework of norm diffusion. In addition, it relies upon the understanding of youth as a general category, thus obscuring other possible identity markers such as gender and class, which might have repercussions on youth action. With such limitations, future research might consider intersectional differences for greater accuracy, other than using a larger, more comprehensive set of data. In conclusion, this paper calls for a more attentive consideration of young people, who should be placed at the center of societal reconstruction.

References

Acharya, A. (2004). How ideas spread: Whose norms matter? Norm localization and institutional change in Asian regionalism. *International Organization, 58*(2), 239–275.

Acharya, A. (2011). Norm subsidiarity and regional orders: Sovereignty, regionalism, and rule-making in the Third World. *International Studies Quarterly, 55*(1), 95–123.

Acharya, A. (2013). The R2P and norm diffusion: Towards a framework of norm circulation. *Global Responsibility to Protect, 5*(4), 466–479. https://doi.org/10.1163/1875984X-00504006

Auerbach, Y. (2004). The role of forgiveness in reconciliation. In Y. Bar-Siman-Tov (Ed.), *From conflict resolution to reconciliation* (pp. 149-176). Oxford University Press.

Baldwin, G. (2019). Constructing identity through commemoration: Kwibuka and the rise of survivor nationalism in post-conflict Rwanda. *The Journal of Modern African Studies, 57*(3), 355–375. https://doi.org/10.1017/S0022278X19000259

Bar-Tal, D. (2000). From intractable conflict through conflict resolution to reconciliation: Psychological analysis. *Political Psychology, 21*(2), 351–365. https://doi.org/10.1111/0162-895X.00192

Bar-Tal, D., & Bennink, G. H. (2004). The nature of reconciliation as an outcome and as a process. In Y. Bar-Siman-Tov (Ed.), *From conflict resolution to reconciliation* (pp. 11-38). Oxford University Press.

Ben-Josef Hirsch, M. (2014). Ideational change and the emergence of the international norm of truth and reconciliation commissions. *European Journal of International Relations, 20*(3), 810–833. https://doi.org/10.1177/1354066113484344

Björkdahl, A. (2006). Promoting norms through peacekeeping: UNPREDEP and conflict prevention. *International Peacekeeping, 13*, 214–228. https://doi.org/10.1080/13533310500437613

Boesenecker, A. P., & Vinjamuri, L. (2011). Lost in ranslation? Civil sciety, faith-based organizations and the negotiation of international norms. *International Journal of Transitional Justice, 5*(3), 345–365. https://doi.org/10.1093/ijtj/ijr018

Brudholm, T., & Rosoux, V. (2009). The unforgiving: reflections on the resistance to forgiveness after atrocity. *Law and Contemporary Problems, 72*(2), 33–49.

Capie, D. (2008). Localization as resistance: The contested diffusion of small arms norms in Southeast Asia. *Security Dialogue, 39*(6), 637–658. https://doi.org/10.1177/0967010608098213

CNBC Africa. (2019, May 6). *The role of art and media in Rwanda's peace and reconciliation efforts* [Video]. YouTube. https://www.youtube.com/watch?v=EqDk3nQPVrc

Clark, J. (2010). National unity and reconciliation in Rwanda: A flawed approach? *Journal of Contemporary African Studies, 28*(2), 137–154. https://doi.org/10.1080/02589001003736793

De Almagro, M. M. (2018). Lost boomerangs, the rebound effect and transnational advocacy networks: A discursive approach to norm diffusion. *Review of International Studies, 44*(4), 672–693. https://doi.org/10.1017/S0260210518000086

De Graaf, A. (2014). Peace like a river: Youth narratives as a current peacebuilding practice. Paper presented at the *International Studies Association (ISA) Annual Convention*, Toronto, Ontario, Canada.

Des Forges, A., & Longman, T. (2004). Legal responses to genocide in Rwanda. In H. M. Weinstein & E. Stover (Eds.), *My neighbor, my enemy: Justice and community in the aftermath of mass atrocity* (pp. 49-68). Cambridge University Press.

Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization, 52*(4), 887–917.

Groß, L. (2015). The journey from global to local: Norm promotion, contestation and localisation in post-war Kosovo. *Journal of International Relations and Development, 18*(3), 311–336. https://doi.org/10.1057/jird.2015.16

iDebate: The power of critical thinking. (2018, November). Queen's Commonwealth Trust. https://www.queenscommonwealthtrust.org/inspiration/idebate-rwanda-critical-thinking-peaceful-debate/

Ikilezi, O. (2019, July 4). 25 years later, a generation has blossomed. *The New Times.* https://www.newtimes.co.rw/news/25-years-later-generation-has-blossomed

Karekezi, U. A., Nshimiyimana, A., & Mutamba, B. (2004). Localizing justice: Gacaca courts in post-genocide Rwanda. In H. M. Weinstein & E. Stover (Eds.), *My neighbor, my enemy: Justice and community in the aftermath of mass atrocity* (pp. 69-84). Cambridge University Press.

Kasherwa, A. C. (2020). The role of youth organizations in peacebuilding in the African Great Lakes Region: A rough transition from local and non-governmental to the national and governmental peacebuilding efforts in Burundi and eastern DRC. *Journal of Peace Education, 17*(2), 123–160. https://doi.org/10.1080/17400201.2019.1688139

Kelman, H. C. (2004). Reconciliation as identity change: A social-psychological perspective. In Y. Bar-Siman-Tov (Ed.), *From conflict resolution to reconciliation* (pp. 111-124). Oxford University Press.

Korman, R. (2014). L'État rwandais et la mémoire du génocide: Commémorer sur les ruines (1994-1996). *Vingtième Siècle. Revue d'histoire, 122*, 87–98.

Kostovicova, D., & Bicquelet, A. (2018). Norm contestation and reconciliation: Evidence from a regional transitional justice process in the Balkans. *Ethnic and Racial Studies, 41*(4), 681–700. https://doi.org/10.1080/01419870.2018.1380211

Kurze, A. (2016). #WarCrimes #PostConflictJustice #Balkans: Youth, performance activism and the politics of memory. *International Journal of Transitional Justice, 10*(3), 451–470. https://doi.org/10.1093/ijtj/ijw014

Lemarchand, R. (2006). Genocide, memory, and ethnic reconciliation in Rwanda. In S. Marysse, F. Reyntjens & S. Vandeginste (Eds.), *L'Afrique des Grands Lacs. Annuaire 2006-2007* (pp. 21-30). Paris: L'Harmattan.

Life & Peace Institute. (2017, September). *"Being and becoming a peacebuilder": Insights from 20,000 hours of youth-led dialogues in the Horn of Africa.* https://life-peace.org/resource/being-and-becoming-a-peacebuilder/

MacKenzie, M., & Sesay, M. (2012). No amnesty from/for the international: The production and promotion of TRCs as an international norm in Sierra Leone. *International Studies Perspectives, 13*(2), 146–163. https://doi.org/10.1111/j.1528-3585.2012.00461.x

McEvoy-Levy, S. (2006). Introduction: Youth and the post-accord environment. In S. McEvoy-Levy (Ed.), *Troublemakers or peacemakers? Youth and post-accord peace building* (pp. 1-26). University of Notre Dame Press.

Mengistu, M. M. (2016). The quest for youth inclusion in the African politics: Trends, challenges, and prospects. *J Socialomics, 6*(1).

Meyerstein, A. (2007). Between law and culture: Rwanda's Gacaca and postcolonial legality. *Law & Social Inquiry, 32*(2), 467–508. https://doi.org/10.1111/j.1747-4469.2007.00066.x

Mgbako, C. (2005). Ingando solidarity camps: Reconciliation and political indoctrination in post-genocide Rwanda note. *Harvard Human Rights Journal, 18*, 201–224.

Ministry of Youth, Culture and Sports. (2005). *National youth policy.* https://www.youthpolicy.org/nationalyouthpolicies/

Opobo, M. (2019, April 9). G25–Mashirika's new youth-produced play for Kwibuka25. *The New Times.* https://www.newtimes.co.rw/entertainment/g25-mashirikas-new-youth-produced-play-kwibuka25#:~:text=%27Generation%2025%27%20(G25),1994%20Genocide%20against%20the%20Tutsi.

Ottendörfer, E. (2013). Contesting international norms of transitional justice: The case of Timor Leste. *International Journal of Conflict and Violence (IJCV), 7*(1), 23–35. https://doi.org/10.4119/ijcv-2948

Peace and Love Proclaimers. (n.d.). *Walk to remember.* Peace and Love Proclaimers. http://peaceloveproclaimers.org/?page_id=99

Peace Direct. (2019, June). *Youth and peacebuilding: Key insights and lessons from a global online consultation.* https://reliefweb.int/report/world/youth-and-peacebuilding-key-insights-and-lessons-global-online-consultation-june-2019

Prisca, S., Kandagor, M., & Kiprono, K. (2012). The perspective of youth in peacebuilding, healing and reconciliation in Eldoret- Kenya. *International Journal of Humanities and Social Science, 12*(2), 187-197.

Rettig, M. (2008). Gacaca: Truth, justice, and reconciliation in postconflict Rwanda? *African Studies Review, 51*(3), 25–50.

Rosoux, V. (2009). Reconciliation as a peace-building process: Scope and limits. In J. Bercovitch, V. Kremenyuk, & I. W. Zartman (Eds.), *The SAGE handbook of conflict resolution* (pp. 543-563). SAGE Publications Ltd.

Rosoux, V. (2015). Post-conflict reconciliation: A humanitarian myth? In P. Gibbons & H.-J. Heintze (Eds.), *The humanitarian challenge: 20 Years European network on humanitarian action (NOHA)* (pp. 47–69). Springer International Publishing. https://doi.org/10.1007/978-3-319-13470-3_4

Rosoux, V. (2017). Reconciliation as a puzzle: Walking among definitions. In V. Rosoux & M. Anstey (Eds.), *Negotiating reconciliation in peacemaking: Quandaries of relationship building* (pp. 15–26). Springer International Publishing. https://doi.org/10.1007/978-3-319-62674-1_2

Russell, S. G. (2015). Global civil society and education policy in post-genocide Rwanda. *International Sociology, 30*(6), 599–618. https://doi.org/10.1177/0268580915605652

Russell, S. G. (2018). Global discourses and local practices: Teaching citizenship and human rights in postgenocide Rwanda. *Comparative Education Review, 62*(3), 385–408. https://doi.org/10.1086/698305

Russell, S. G. (2019). *Becoming Rwandan: Education, reconciliation, and the making of a post-genocide citizen.*

Rutgers University Press.

Rwanda: Walk to remember. (2018, April 25). *The Independent.* https://www.independent.co.ug/rwanda-walk-to-remember/

Salomon, G. (2002). The nature of peace education: Not all programs are created equal. In G. Salomon & B. Nevo (Eds.), *Peace education: The concept, principles, and practices around the world* (pp. 3-14). Lawrence Erlbaum Associates, Publishers.

Saunders, H. H. (2011). *Sustained dialogue.* Civicus.org. https://www.civicus.org/documents/toolkits/PGX_D_Sustained%20Dialogue.pdf

Shipler, M. (2008). Nepal's youth as peacebuilders. *Peace Magazine, 24*(2), 8-10.

Sikkink, K., & Kim, H. J. (2013). The justice cascade: The origins and effectiveness of prosecutions of human rights violations. *Annual Review of Law and Social Science, 9*(1), 269–285. https://doi.org/10.1146/annurev-lawsocsci-102612-133956

Staub, E. (2013). Building a peaceful society: Origins, prevention, and reconciliation after genocide and other group violence. *The American Psychologist, 68*(7), 576–589. https://doi.org/10.1037/a0032045

Stimmer, A. (2019). Beyond internalization: Alternate endings of the norm life cycle. *International Studies Quarterly, 63*(2), 270–280. https://doi.org/10.1093/isq/sqz001

Subotic, J. (2015). Truth, justice, and reconciliation on the ground: Normative divergence in the Western Balkans. *Journal of International Relations and Development, 18*(3), 361–382. https://doi.org/10.1057/jird.2015.13

Thapa, M. (2009). From conflict to peace: the role of young people in recreating Nepal. In Corcoran, P. B. & Osano, P. M. (Eds.), *Young people, education, and sustainable development: Exploring the principles, perspectives and praxis* (pp. 289-295). Wageningen Academic Publishers.

Tholens, S., & Groß, L. (2015). Diffusion, contestation and localisation in post-war states: 20 years of Western Balkans reconstruction. *Journal of International Relations and Development, 18*(3), 249–264. https://doi.org/10.1057/jird.2015.21

Tiemessen, A. E. (2004). After Arusha: Gacaca justice in post-genocide Rwanda. *African Studies Quarterly, 8*(1), 57–77.

UNESCO. (2013, May). *Promoting a culture of peace and non-violence in Africa through education for peace and conflict prevention.* http://education4resilience.iiep.unesco.org/en/node/818

Unoy – Global network strengthening sustainable youth-driven peacebuilding. (n.d.). Unoy - Global Network Strengthening Sustainable Youth-Driven Peacebuilding. https://unoy.org/

UNSCR 2250: Introduction. (2015). Youth4Peace Portal. https://www.youth4peace.info/UNSCR2250/Introduction

Vuković, S. (2020). Peace mediators as norm entrepreneurs: The EU's norm diffusion strategy in Montenegro's referendum on independence. *Swiss Political Science Review, 26*(4), 449–467. https://doi.org/10.1111/spsr.12424

Weinstein, H. M. & Stover, E. (2004). Introduction: conflict, justice and reclamation. In H. M. Weinstein & E. Stover (Eds.), *My neighbor, my enemy: justice and community in the aftermath of mass atrocity* (pp. 1-26). Cambridge University Press.

Wienand, C. (2012). From atonement to peace?: Aktion Sühnezeichen, German-Israeli relations and the role of youth in reconciliation discourse and practice. In Schwelling B. (Ed.), *Reconciliation, civil society, and the politics of memory: Transnational initiatives in the 20th and 21st century* (pp. 201-236). Bielefeld: Transcript Verlag.

Wiener, A. (2004). Contested compliance: Interventions on the normative structure of world politics. *European Journal of International Relations, 10*(2), 189–234. https://doi.org/10.1177/1354066104042934

Wiener, A. (2017). A theory of contestation—A concise summary of its argument and concepts. *Polity, 49*(1), 109–125. https://doi.org/10.1086/690100

Williams, P. D. (2009). The Responsibility to Protect, norm localisation, and African international society. *Global Responsibility to Protect, 1*, 392.

Zorbas, E. (2009). What does reconciliation after genocide mean? Public transcripts and hidden transcripts in post-genocide Rwanda. *Journal of Genocide Research, 11*(1), 127–147. https://doi.org/10.1080/14623520802703707

Social Sciences

Social Media Platforms and a New Regime of (In)visibility

Exploring the Socio-Political Implications of Algorithmic Power

Salomé Petit-Siemens

Supervisor
Stacey Vorster (UvA)
Reader
Dr. Anne de Graaf (AUC)

Photographer: Richard Essink

Abstract

The suspension of Donald Trump's Twitter and Facebook accounts after the violence on Capitol Hill has rekindled discussion on the expansive social role of internet platforms and their control over public discourse online. Social media companies are designing networks, powered by algorithms, which can sort, rank, and censure human experience and people. This suggests that networked experience is being configured through visibilities. In light of this phenomenon, this thesis conceptualises a new regime characterised by the management of visibility through algorithmic systems. So as to explore the broader significance of the regime, the concept of (in)visibility is introduced to interrogate questions of power and, ultimately, the socio-political implications of algorithms. The term is theorised through Foucauldian themes of power, subjection, and resistance. Methodologically, several case studies are examined to explore a set of questions: How do algorithms, as techniques or forms of power, become means to produce, order and manage visibilities? "How," not in the sense of "How does power manifest itself?" but "By what means is it exercised through social media architecture?" and "What happens when individuals choose to resist it?" In essence, this paper explores how the concentration of power, enabled by social media architecture, is giving rise to symbolic forms of control through (in)visibilisation.

Keywords: *(In)visibility, algorithms, power, social media, resistance*

Suggested citation:
Petit-Siemens, S. (2021). Social media platforms and a new regime of (in)visibility: Exploring the socio-political implications of algorithmic power. *AUC Undergraduate Journal of Liberal Arts and Sciences, 16*(1), 69-95.

Contents

List of Figures and Tables

1 Introduction

The Covid-19 pandemic has both revealed and fuelled our dependency on online platforms, generating an unprecedented semblance of hyperconnectivity. There lies great power in the design of the communication technologies that constitute a large part of our daily lives. Large companies such as Facebook and Twitter are single-handedly controlling the architecture of vast digital spaces, deciding on the rules to access them as well as choreographing the social interactions they host. Consequently, they confer the possibilities of visibility for billions of users; first by determining what users can see in terms of content and who they can see in terms of other users, and second, by deciding on the platforms' configurations, namely how the possibilities of visibility are conferred. Since social media platforms have become the milieu where an important part of our experience with political life unfolds (Beer, 2017), concerns have arisen over the expansive social role of internet platforms and their control over public discourse online (Zuboff, 2021).

Notably, the dawn of internet platforms has transformed the pace and scale of online interaction. With it, the architecture of platforms has evolved to accommodate growing interactivity. In this web of interactivity, users and their subjectivity are turned into nodes: end objects, usually represented as circles in network maps, that amass information about the individual user and their interaction pathways. Social media platforms like Facebook, Instagram or Twitter consist of webs of nodes, or users, connected by "vectors"—relational lines with dialectical flows of visibility through which nodes have, in principle, the capacity of seeing one another (Brighenti, 2010, p. 45). Yet, based on profit-driven logics of engagement, companies have instituted algorithmic systems into the design of platforms in order to sort, rank, and censure content (Gillepsie, 2014). Through these mechanisms, certain flows become hypervisibilised while others are invisibilised and restricted. Take, for instance, the events leading up to and after the Capitol Hill incident on January 6th, 2021. Inflammatory tweets by former United States president, Donald Trump, on electoral fraud, were amplified and made hypervisible through high engagement—retweets, comments, reactions. However, in the aftermath of the Capitol attack, Trump's account was suspended for incitement of violence and his presence was thereby rendered invisible on several platforms. As this case illustrates, algorithms do not solely produce and order networks of nodes and vectors, but actually carry political implications whereby human experience and people are subjected to processes of visibility and invisibility through platforms' shifting socio-technical norms (Brighenti, 2010). This suggests that networked experience is ultimately managed through (in)visibility.

In line with theories on media power, critically engaging with questions of (in)visibility—whom/what should see whom/what—resonates with theories on visibility regimes (Brighenti, 2010).[1] As conceptualised within media studies, regimes of visibility are patterns of interaction that "establish orders and modes of seeing and being seen, of exposure and concealment or invisibility" and which inevitably elicit mechanisms of power (Krasmann, 2017, p. 11). This is not to say that algorithmic systems are agencies of power, but rather that they can be thought of as techniques enacting a new form of visibility regime online.[2] Through this configuration of visibility, the socio-political implications of social media algorithms can be questioned, and their power effects examined. These effects are not so much concerned with visibility solely in terms of seeing but rather with the question of visibility with regards to its connection to knowledge and power—the knowable and the unknown—, a rapprochement to Jacques Rancière's conceptualisation of the "distribution of the sensible" (2013, p. 12). Following a similar line of thought, scholars examining algorithmic control are primarily concerned with the processes through which our social reality is curated online (Beer, 2017; Bucher, 2012; Cristiano & Distretti, 2020; Musiani, 2013; Zuboff, 2021). Dynamics of data extraction enable algorithms to order content according to patterns of user consumption (Cristiano & Distretti, 2020). Although this extraction becomes "willing" through the voluntary participation of users with the platform, algorithmic mechanisms remain largely opaque (Beer, 2017), creating visibility asymmetries between user and platform. The process of data extraction coupled with visibility management on social media seems to point to a decline in individual autonomy online, a concept that underlines the main theoretical discussions on resistance in the field of critical media studies (Beer, 2017). By exploring strategies of resistance against algorithms, this thesis aspires to counteract the asymmetries of power by unveiling the ways in which, mediated by technologies, knowledge circulates under a veil of invisibility.

While the concept of visibility is explicitly employed or alluded to in critical media scholarship, it has rarely been theoretically conceptualised with regards to algorithmic control. Therefore, a new understanding is necessary, especially in the context of the Covid-19 pandemic where most daily activities take place online. This research departs from previous conceptualisations focused on media theory, and instead grounds the algorithmic management of visibility in its socio-political ramifications. Through the concept of (in)visibility, this thesis offers a theory for a new form of visibility regime in social media platforms. It does this by dissecting Foucault's (1982) theory of subjection, power, and control, but

[1] "The styles of visibilisation thus correspond to certain visibility regimes, which are constitutive of the domain of the public, and how bodies enter this domain" (Brighenti, 2010, p. 52). Media and communication studies have examined regimes of invisibility in the past. In broadcasting and other traditional media, media outlets function as the gatekeepers of information where "many see one" (Brighenti, 2010, p.52). This is different from new forms of media in the information age where "many can see many."

[2] I refer to techniques because they are not agential but rather semi-automated, ultimately created and tweaked by human beings. As such, they become techniques through which to exercise power. This will be further conceptualised in Chapter 1.

also of resistance, to answer the following questions: How do algorithms, as techniques or forms of power, become means to produce, order and manage visibilities? "How" is not employed in the sense of "How does power manifest itself?" but rather "By what means is it exercised through social media architecture?" and "What happens when individuals choose to resist it?"

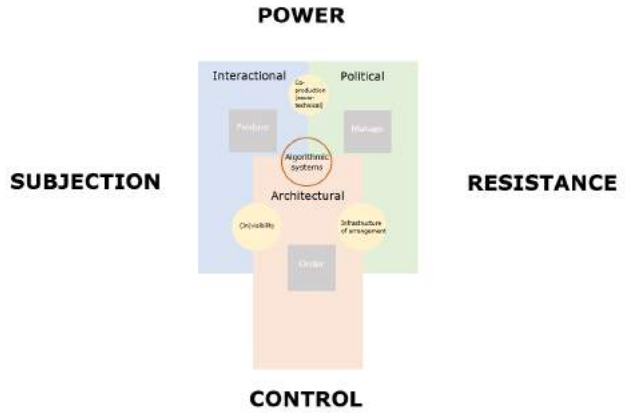

Figure 1: Sketched diagram for the theoretical trajectory proposed in this study. It combines and relates the main thematics of the paper together. The heading of power relates to the concept of control and in the same manner, subjection and resistance are two faces of the same coin— resistance already exists as a facet of subjection. Yet, the diagram can also be read in circular motion: Power-Subjection-Control-Resistance.

To answer these questions, this study devises a tripartite regime of visibility (see Figure 1). The online regime of visibility is constituted and enabled by the architectural and interactional aspects of social media platforms. Furthermore, this research takes on another mode of studying regimes of visibility, as proposed by Foucault, through an empirical investigation of power relations and strategies of resistance. The structure of the paper follows the same logic, taking form in three separate chapters, also in par with the tripartite research question. After laying out the theoretical dimensions of the paper in the research context and methodology section, the remaining part of the paper proceeds as follows:

Concerned with the architectural aspect of the regime, Chapter One focuses on the materiality of social media platforms. Through an analysis of social media network maps, patterns of (in)visibility become apparent, shedding light on the implicit rules of arrangement within platforms. These rules of arrangement, implemented by algorithmic work, confer the possibilities of visibility for billions of users. Ultimately, it is argued, the creation of behavioural patterns and subsequent ordering of reality, create systems of governance that ultimately shape social order, one of the main preoccupations of this study at large. In this way, algorithms, as techniques of power, are first politicised in this chapter through how they order users and content online in mechanical ways.

Despite the emphasis on the material politics of algorithms in Chapter One, materiality cannot be detached from the domain of the social. This is because algorithms function in socio-technical assemblages of people and political economies. Thus, Chapter Two expands on the previous chapter by positioning itself at the intersection of the technique, human bodies and experiences: what happens when human lives are the input and the output? This chapter touches upon political economies and data extraction, further delving into subjection and what it means to be a "good" (data) subject on social media platforms. In addition, it revolves around the concept of production, not only co-production of humans and machines but also production of reality by algorithms and its socio-political implications. Finally, a layer of analysis will be added to the network maps. The semantics of "human encounters with algorithms" will be addressed through an examination of the meaning of "algorithmic imaginaries" as a first step towards algorithmic consciousness and towards resistance.

Finally, through a theoretical engagement with Michel Foucault and case studies on algorithmic resistance, Chapter Three develops the concept of algorithmic resistance. It begins by grouping algorithmic strategies of resistance in active and non-compliant categories and then delves into examples. The types of active resistance include semiotic resistance, which involves changing words to acronyms or types of code language to avoid censure, hashtag takeovers, redeployment of algorithms for antagonist use and lastly, instances of non-compliant resistance include disconnection or obfuscation.

This thesis aims to contribute to a deeper understanding and demystification of how algorithms on social media work. The socio-political grounding of this research can yield greater awareness of the exploitative dynamics of social media and take part in the process of demanding more transparency and accountability from these platforms.

2 Research Context

A large and growing body of literature in social science research has examined the expression of power in the context of algorithmic work (Beer, 2017; Brunton & Nissenbaum, 2015; Bucher, 2017; Cristiano & Distretti, 2020, Musiani, 2013; Nissembaum, 2011). Among different lines of scholarly inquiry, three salient themes have dominated the conversation. First, the theme of online visibility is an overarching scholarly preoccupation. Literature on visibility is most concerned with exploring how algorithms manage visibility online and consequently assemble online social reality. Second, literature on governance expands on the subject of visibility by adding themes of data extraction and profit maximisation by social media companies, which further problematise questions of transparency and the power of algorithms to

order social reality online (Beer, 2009; Beer, 2017). The third theme ties into resistance, which examines different antagonisms individuals develop to challenge algorithmic power (Brunton & Nissenbaum, 2015; Ettlingler, 2018; Treré et al., 2020; Velkova & Kaun, 2021). By examining power through these themes, the social and political implications of social media algorithms are being written out. This paper mirrors and expands on these themes in the subsequent chapters.

Drawing on visibility as the central tenet to their research, Cristiano and Distretti (2020) have scrutinised the ability of algorithms, as semi-autonomous agents,[3] to manage visualities. They argue that algorithms act as mediators of reality, in essence managing what is and is not seen online (Cristiano & Distretti, 2020). Backed by underlying themes of data extraction and the profit-driven nature of social media platforms, they illustrate the power asymmetries embedded in networks (Cristiano & Distretti, 2020). Algorithms function as invisible mechanisms that render users' data hypervisible without their full knowledge (Cristiano & Distretti, 2020). This issue underlines a wider preoccupation in the field, concerning the ways in which algorithms function under a veil of invisibility. Some writers have propounded theories on this concern. For instance, Thrift (2004) put forth the concept of the "technological unconscious" (p. 213)—the operation of invisible and powerful information technologies that come to produce everyday life. In recent years, a broader perspective has been adopted by Pasquale (2015). In his book The Black Box Society, he explains how the internal workings of the algorithms that pervade a great part of our lives, remain hidden for commercial purposes. Precisely, algorithms create power asymmetries between the technology and the user, manifested in the opacity of algorithmic work but also in the ways algorithms manage visibility. Expanding on this, Musiani (2013) contends that algorithms produce political and social hierarchies online by prioritising topics of discussion as well as certain users over others. Undoubtedly, hierarchies of content also create asymmetries between users, an issue that has generated scholarly discussion around visibility and recognition (Brighenti, 2010; Cristiano & Distretti, 2020; Zuboff, 2021). Brighenti (2010), in particular, reflects upon the consequences of visibility for the public sphere and digital citizenship. If political visibility and recognition is at play on social media platforms, the implications of algorithms go well beyond the ordering of visualities.

Visibility asymmetries become more problematic when considering that most of our political experience online is being mediated by algorithms (Beer, 2017). Consequently, scholars like Beer (2017) and Just and Latzer (2017) have chosen to conceptualise the power of algorithms through the theme of governance. For one, Beer (2017) is concerned with how "software may be taking on some constitutive or performative role in ordering that world on our behalf" (p. 4) and ultimately holding the function of constructing reality online. This reasoning resonates within established theoretical frameworks on social reality. For example, Berger and Luckmann (1967) theorise that social order is based on a shared social reality that emerges from real or symbolic interactions. In this sense, the online reality produced by algorithmic work can certainly influence behaviour and action in everyday life. Indeed, software shaping the everyday experiences of users becomes an expression of power, not only over us, but fed and reproduced by human engagement online (Beer, 2009). Precisely, the naturalisation of data extraction, the creation of behavioural patterns, and the ordering of reality, create systems of governance that ultimately shape social order (Cristiano & Distretti, 2020). Most concerning is the fact that algorithms are designed for private companies' profit maximisation, which, coupled with the lack of transparency, makes algorithms hardly fit to govern spaces of social interaction (Just & Latzer, 2017). As noted by Just and Latzer (2017), the decision-making powers of technologies generate questions concerning human agency on social media platforms. For instance, should the public have a say in how platforms are governed?

Another field that tackles the technological challenges to human agency focuses on the notion of resistance. Scholarly methods for analyzing expressions of algorithmic power do not only involve visibility and governance but also an examination of antagonisms: individual and collective instances of algorithmic resistance. Neumayer et al. (2021), for example, suggest that online visibility wavers between empowering and disempowering poles. While some people, for instance activists, might manipulate algorithms to become more visible, other, more vulnerable, groups might seek to resist algorithmic visibility. Either way, resistance seems to constitute a zero-sum game. The mere interaction with social media platforms creates a dynamic of human-machine coproduction and therefore a certain user compliance to data extraction (Cristiano & Distretti, 2020). As such, disconnection, anonymity, and invisibility, are budding ways to opt out of algorithmic power (Treré et al., 2020). Yet, most users still choose complicit ways of resistance—resisting while staying connected—because the offline state can lead to consequences such as loss of contact or job prospects.

Nevertheless, resistance also presents an empowering potential and can uncover, through the strategies that individuals use to resist them, the workings of algorithms. According to Vickers and McDowell (2021), the idea of resistance to algorithms also reveals an important paradox: social media companies have mechanisms in place that function on exclusion, but making platforms more inclusive might be detrimental to those who wish to remain invisible to algorithmic power. This paradox demonstrates that visibility regimes are in fact controlled by technologies of power, which can bring about socio-political consequences for users.

By mirroring and expanding on these themes in the rest

[3] I refer to them as semi-autonomous agents because they function with little to no human intervention apart from the initial coding and periodic tweaks to ameliorate effectiveness.

of this paper, I merge adjacent fields of media and algorithmic studies, philosophy, and science and technology studies to contribute to the academic objective of bringing more visibility to the "black box" conundrum (Pasquale, 2015), essentially to uncover the technical and socio-political workings of algorithms.

3 Methodology

There are a number of instruments available to investigate the dynamic of power present in algorithmic functioning. Recent social science research has examined algorithms in an isolated manner, that is, studying algorithms as single artefacts (i.e., examining the code, reverse engineering, investigating input and output). However, because algorithms function in highly dynamic platforms, solely unveiling algorithm code would fail to account for the relational, contingent, and contextual environment in which they function (Kitchin, 2017). Indeed, the limited knowledge that exists on social media algorithms often relegates the socio-political realities of these technologies to a second plane, disregarding how algorithmic systems in social media are entangled with human experience and political economies (Kitchin, 2017). As Foucault (1978) contends in his lecture The Birth of Biopolitics, the role of social research is "not to discover what is hidden but to render visible what precisely is already visible, so intimately linked to ourselves, that as a consequence, we don't see it" (p. 540). Therefore, this paper re-conceptualises algorithms within the wider context of their socio-technical assemblages, referring to social media as polycentric environments in which linkages and mediations occur in a milieu of human experience, political economies, and technologies of power.

The concept of (in)visibility, central to this work, conjoins the sensorial field of the visual and the broader socio-technical field of visibility. This is to say, the term is able to interlink algorithmic mechanisms of ordering and ranking to its larger significance in terms of knowledge and, ultimately, power. In particular, the concept of (in)visibility captures the ordered but contingent and opaque effects of algorithmic visibility. I propose this concept as open-ended, an element in which the social occurs, and through which to analyse the complicated interplays between autonomy/subjection and power/resistance embedded in algorithmic work. In this way, the concept aims to capture the ambivalent experience of the user with social media platforms.To be clear, these dyads are not mutually exclusive but rather coexist and interact, producing visibility asymmetries. Once organised in regimes, visibility asymmetries reproduce the power/knowledge asymmetries of which they are a product (Brighenti, 2010). For instance, the algorithms can create self-perpetuating loops whereby users who are less visible are less often objects of interaction, therefore, further decreasing their possibilities of visibility. These cyclical and interrelated dynamics characterise our interpretation of a new mode of visibility regime online, which, enabled by the architectural and interactional character of social media platforms, exerts political influence over the way we interact online. This is not to say that social phenomena online can be interpreted solely in terms of visibility. Rather, this paper utilises visibility as an overarching analytical tool to understand new technologies of power. For this, new diagrams of power are needed to trace how they operate in new architectures like social media. As a novel mode of analysis, this paper proposes the examination of network maps, which are important tools that can help us visualise patterns of (in)visibility. Network maps are able to show differences in interactive density or interconnectivity and, therefore, delineate the implicit rules of arrangement within platforms.

According to Foucault (1982), the study of power conflates to analysing the new modes through which humans are made subjects; through these processes of subjectification, one can uncover power relations. In the context of algorithmic control, the mode of subjection is enabled through rendering things visible (Krasmann, 2017). In effect, on social media, there only exists what has been ordered and rendered visible by algorithms. Thereby, the user, as a subject of power, is constituted online by the algorithmic mechanisms of seeing and knowing (Krasmann, 2017). What is paramount here is to interrogate the ways in which knowledge circulates and functions—by sorting, ranking, and censuring—and how these functions relate to power. Indeed, the question of power concerns how rather than what is made visible, and through which specific infrastructures of arrangement. Following this line of thought, Foucault sought to explain the regulatory force of power inherent in specific architectural configurations. As mentioned, the architecture of social media encompasses nodes and vectors that do not merely create a network but have socio-political implications, where people and their experiences are made visible or invisible. Similar to other modes of classification infrastructures that have been scrutinised in the past (e.g., bureaucracy), algorithms have demonstrated the ability to assemble social and material worlds online.[4] Undoubtedly, the management of visibility online involves a certain power over the representation of individual online realities for billions of people (Zuboff, 2021). As explained by Berger and Luckmann (1967), this has consequences on a shared social reality and consequently, the formation of social order.

Mechanisms of subjection cannot be examined without interrogating their relation to mechanisms of exploitation and domination, as all forms interact in complex and interlinked ways (Foucault, 1982). The naturalisation of data extraction and commodification of human experience on social media platforms follows logics of exploitation. The voluntary partic-

[4] A "world" in this sense is also very much in line with Jacques Rancière's conceptualisation of the "distribution of the sensible" (2013, p. 12). Although beyond the scope of this paper, this analysis could be enriched by engaging with Rancière since many of the arguments overlap with his theory.

ipation of users with Internet platforms might indicate prior consent, but as Foucault (1982) contends, power relations are not by nature a product of consensus— this would assume a passive subject. Yet, in his revision of his prior text on discipline (Foucault, 1995), Foucault regards the subject as active (Foucault, 1982). In fact, Foucault declares that action is the very condition that enables subjection, which is ultimately the basis of the power relation. In this sense, the voluntary engagement of users with platforms actively generates data, input without which algorithms cannot function. In this human-machine co-production, users contribute to their own subjection.

The new understanding of subjects as active agents in The Subject and Power allows for new conceptualisations of power relations (Foucault, 1982). Foucault (1982) suggests investigating forms of resistance as a starting point of analysis. As such, to understand algorithmic power, this paper lays the theoretical basis in the first two chapters and then, contextualises it in the last chapter, through case studies of active and passive algorithmic resistance. Therefore, this paper engages with the Foucauldian thematics present in algorithmic work: subjection, power and control, but also resistance, to answer the following questions: How do algorithms, as techniques or forms of power, become means to produce, order and manage visibilities? "How," not in the sense of "How does power manifest itself?" but "By what means is it exercised through social media architecture?" and "What happens when individuals choose to resist it?"

4 Chapter 1: Architecture, (In)visibility, and Technique

To characterise the interpretation of a new online regime of visibility and its socio-political influence, this chapter investigates the architecture mediating the visibilities encountered online. While a detailed technical explanation of algorithms seems unnecessary, this work engages the logics of social media architecture, illustrated in this chapter by network maps, in order to lay the grounds for an informed theoretical discussion. Accordingly, the emphasis of this chapter is the materiality of online socio-technical assemblages, thereby setting out to explore some of the many ways in which algorithms express power. By questioning how online visibility is constrained or enabled—and through which arrangements, techniques and modes of governance—we can begin to understand an online regime of visibility that unfolds through social media architecture, and moreover, further develop the concept of (in)visibility.

Social media architecture is a composite of nodes or users, connected through relational and dialectical lines through which everyone has, in principle, the capacity of intervisibility—when users are mutually visible to one another (Brighenti, 2010). As it hosts social interaction online, the platform becomes a site of circulation of vectors in different directions, which are defined both top-down, through platform norms, and bottom-up, shaped by user interaction. Platforms are decentralised and ontogenic digital infrastructures that "facilitate and shape personalised interactions, organised through the systematic collection, processing, monetisation, and circulation of data" (Poell et al., 2019, p. 3). This data is managed through algorithmic sorting, prioritising certain metadata over others, hence shaping what information and which users become visible, and therefore deciding what is left unseen. These important processes of prioritisation and marginalisation are often invisible, relegating the definition of algorithms to a broad understanding of them as "encoded procedures for transforming input data into a desired output, based on specified calculations" (Gillespie, 2014, p. 167). A limited understanding of the workings of algorithms, makes understanding how they confer visibility complicated: under what conditions is a user made visible? Answering this question requires examining what input is fed into the algorithm, what meta tags (likes, retweets, mentions, etc.) are drawn upon to make connections between users, and how personalisation determines what individuals encounter online. Diagrammatics of algorithms, like the ones in Figures 2, 3 and 4, are useful tools for understanding these processes, as they portray user interactivity around a topic. In this respect, the visualisation of big data offers a tangible dimension where nodes and edges do not simply create a network but represent politics, in which users are deliberately made visible or invisible.

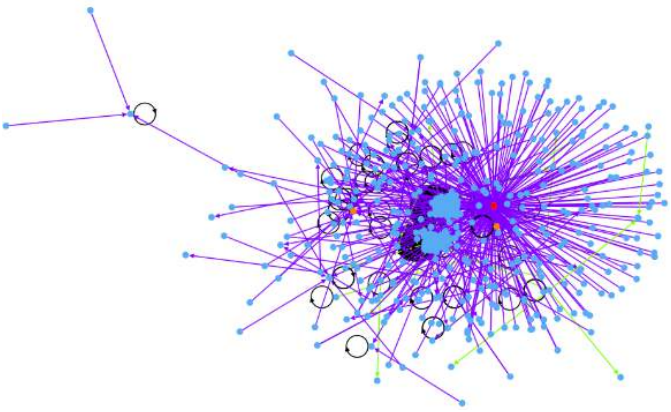

Figure 2: Network map without interconnectivity. Created with NodeXL using Twitter search with terms "#RIPTwitter" and "algorithm." The blue nodes represent users. Nodes in orange or red amass the highest interaction and are therefore the most visible users in the network.

Figures 2, 3 and, 4 are network maps that represent interaction on Twitter surrounding the hashtag #RIPTwitter. #RIPTwitter was the backlash response of users to a BuzzFeed article that claimed Twitter was going to modify the organisation of its timeline from simple chronological order (new to old) to an algorithmically determined one, based

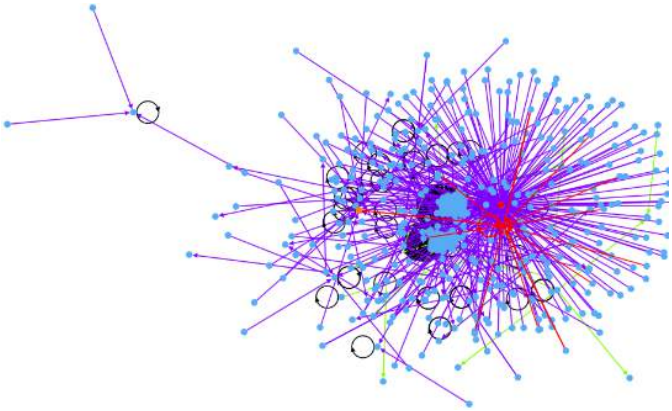

Figure 3: Network map with medium interconnectivity. The nodes towards the periphery of the map have little to no interaction with the users at the centre of the conversation.

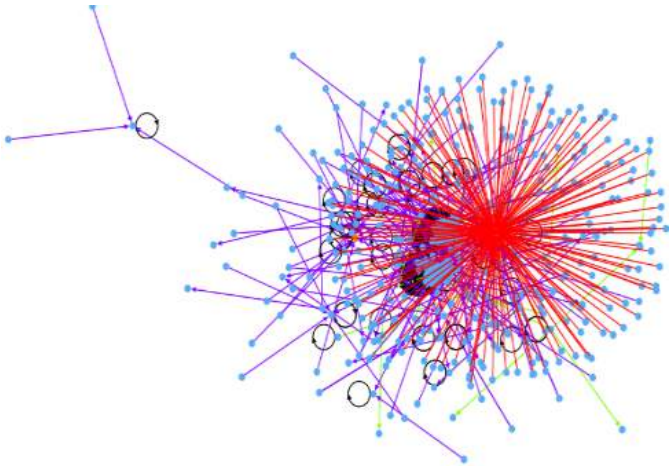

Figure 4: Network map with high interconnectivity. The node at the centre of the red interaction has the highest visibility in the network and is usually the main interlocutor in the conversation.

on relevance to each user (Kantrowitz, 2016). This change implied that users would no longer have the autonomy to choose what information was relevant to them since the algorithm would do that for them, exposing users only to the content established—by unknown parameters—as relevant. The news prompted a reaction on Twitter, with users posting the hashtag #RIPTwitter as a form of protest. While the semantics of the movement—that is, the language of the tweets—is revealing in itself, this stage of the analysis focuses solely on the materiality of the network maps. The map shown above was created by searching the terms "#RipTwitter" and "algorithm." Orange and red nodes signal the most popular users in the conversation, while black circles represent single tweets. Mentions,[5] which constitute a great part of the activity in these maps, are depicted in purple. By clicking on different nodes, one can distinguish variations in user visibility evidenced by the relational vectors that light up in red. Evidently, the maps show differences in interactive density and interconnectivity. For instance, Figure 4 details a cluster around a certain user. This concentration signals a key user, often a person leading the conversation. The content these users create and share is often the most popular and widely reproduced in these networks. In contrast, nodes towards the periphery of the network and black circles reflect relatively unconnected Twitter users who tweet about a topic but hardly interact with others in the discussion.

The differences in behaviour patterns and interaction are not arbitrary. They elicit a series of normative questions about modes of governance online, namely whom/what should see whom/what. In fact, asking what algorithms reflect by observing the output of an algorithmic process provides an inkling about their functioning. This involves denying and allowing, incentivising and disincentivizing possibilities of visibility (what can or cannot see or be seen) and configurations (how something or someone can or cannot see or be seen) (Brighenti, 2010). For instance, inscribed in Facebook's EdgeRank[6] and Twitter's Trending Topics[7] are algorithms that prioritise some topics and people over others (Bucher, 2018). Facebook's algorithm is automatically inclined towards emotive and reactive answers to a subject (Bucher, 2018). Along similar lines, Twitter's Trending Topics feature enables users to boost certain topics by consistently tweeting about a specific event, while retweets and mentions can heighten the visibility of a specific tweet. However, Twitter also actively pushes promoted tweets—paid for by companies and personalised via algorithms to fit specific user profiles (Bucher, 2012; Dijck & Poell, 2013). This type of interaction is paramount. While some users become visible through modes of reception (i.e., likes, shares, posts, retweets), other not so active profiles fade into the background (Brighenti, 2007; Neumayer et al., 2021). Though internet platforms were initially designed as democratic spaces of participation (Web 2.0: see Beer, 2009), the workings of algorithms suggest that desirability can be mapped in the specific mechanisms of visibility (Bucher, 2018). In this way, the diagrammatics of algorithms exemplify patterns of inclusion and exclusion, partly revealing the metrics that determine what is not only relevant, but also appropriate and legitimate.

A close examination of Facebook's algorithm that censors content and polices political expression in 2017, revealed the favouring of elites and governments over activists and racial minorities (Angwin, 2017). Based on principles of equality,

[5] A mention is a type of interaction between two users in Twitter where one user uploads a post containing another account's Twitter username, preceded by the "@" symbol. This type of interaction usually involves responding to the other user's post.

[6] EdgeRank is an algorithm that is embedded in Facebook and is utilised to order the News Feed on the platform based on what type of content the user interacts with (e.g., liberal or conservative content), how they engage with it (through which functions e.g., likes, comments, mentions, etc), and the newness of the content. For more information on how the algorithm functions, see Bucher (2012).

[7] According to Twitter's FAQs, the trending topics that appear on users' feed "are determined by an algorithm and, by default, are tailored for you based on who you follow, your interests, and your location. This algorithm identifies topics that are popular now, rather than topics that have been popular for a while or on a daily basis, to help you discover the hottest emerging topics of discussion on Twitter."

the algorithm was trained to censure posts if they mentioned "protected categories"—based on race, sex, gender identity, religious affiliation, national origin, ethnicity, sexual orientation, and serious disability/disease (Angwin, 2017). Yet, posts containing composite categories, for instance "black children" or "female drivers" were recognized by the algorithm as appropriate (Angwin, 2017) (Figure 5).

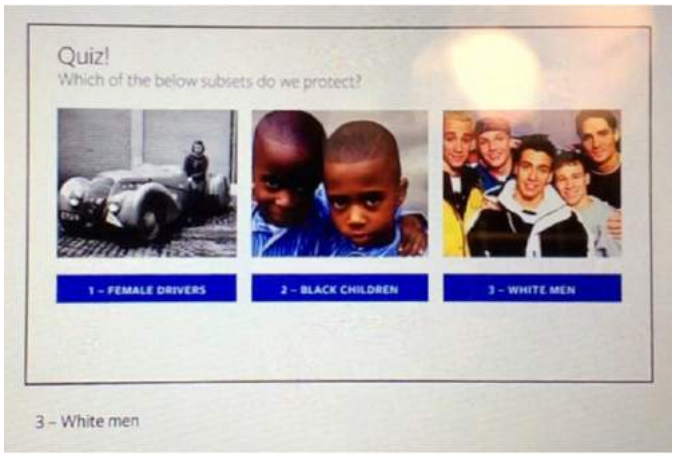

Figure 5: Screen used to train the hate speech algorithm. Facebook trains algorithms to censor and delete hate speech against 'protected categories' such as white males, but allow attacks on 'subsets' such as female drivers and black children. Source: ProPublica.

Consequently, posts containing categories such as "white men" were being taken down, while expressions like "radicalized Muslims" were left free to roam the digital space. The fact that "white men" is considered a protected category by these algorithms reveals that the regime of visibility online is reflective of societal biases and inequalities. Essentially, the algorithmic rules hold the power to differentiate between hate speech and legitimate political expression online. Similarly, Facebook and Twitter also introduced a measurement of legitimacy on their platforms, with fact-checking labels appearing next to posts during the presidential elections, and more recently during the pandemic (Figure 6). However, though these mechanisms arbitrate relationships between users and content around contentious public issues, the standards for appropriateness or legitimacy of content remain ambiguous.

In this way, by managing the visibility of users and topics, algorithms can encourage and discourage discussion while setting the political and social agenda as well as designating the interlocutors that are central to the conversation (Musiani, 2013). In essence, algorithms structure behaviour by enabling and constraining decisions, and by making certain users or topics visible and thus, knowable. However, building on the Foucauldian sensitivity for the "microphysics" of power (Foucault 1979: 92-102), this study reinstates that the question of algorithmic power ultimately concerns how rather than what is made visible. This requires dismantling the inner workings of algorithms as techniques that put the very power into effect. This research broadly defines algorithms as techniques instead of the standard science and technology studies technical definition for two reasons. Firstly, techniques are essential to architectural compositions; they are complicit in the mechanic and calculated modes of ordering multiplicities that sustain architectures of power (e.g., discipline as the technique of the Panopticon). Secondly, through the concept of technique, we can aggregate the themes of capitalist extraction, deregulation, and standardised means of governance, which are present in the social media regime of visibility.

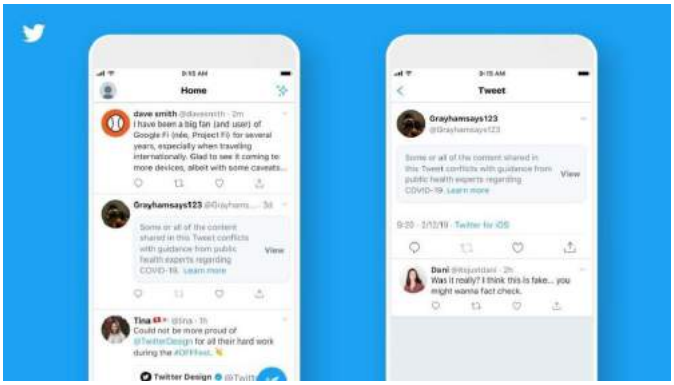

Figure 6: Covid-19 fact-checking labels on Twitter. Source: Google.

According to Ellul (2006), technique refers to "any complex of standardized means for attaining a predetermined result" (p. 7).[8] Theoretically stated, it converts spontaneous and unreflective behaviour into behaviour that is deliberate and rationalised through a set of norms or instructions (Ellul, 2006). Algorithms, as techniques, establish what Lash (2007) considers "algorithmic generative rules" (p. 70), a set of standards modelled on societal ideas of efficiency and fairness. Nonetheless, algorithms are not neutral, objective technologies. They are mechanical and procedural insofar as they strive for the best approximation to their predetermined objective (instructions), results that users interpret in a substantive manner (Cardon, 2018). Therefore, these techniques are not powerful in themselves but only to the extent that they are designed and tweaked by humans, thereby reflecting social norms. Notably, although the coding of algorithms demands the rigour and objectivity characteristic of scientific advancement, when applied to social media platforms, the functioning of algorithms seems to move away from those demands.[9] This is exemplified in the previous example where "white men" is categorised in algorithmic training as a "protected category."

[8] Scott Lash (2007) sets forth the moderative and commercial imperatives of algorithms. He explains that the engagement logics of covert algorithms are conduits of capitalist power, making power increasingly embedded in the algorithm. As such, for Lash, power in society becomes ontological, embedded in code, and hence more difficult to resist through traditional means.

[9] Langdon Winner, on the other hand, argues in his article "Do Artifacts have politics?" (1986) that political ideologies are embedded in technologies, making them inherently political.

Furthermore, algorithms express power in how they function under a veil of invisibility and in how they run largely unscrutinised. Indeed, algorithmic work involves the mobilisation of specialised knowledge and the mastering of calculative and communication technologies (Hansen & Flyerbom, 2015). Thus, their understanding is out of reach for most people. As such, techniques of classification and their crucial role in the building of information infrastructures have been core preoccupations of science, technology, and social science scholars for several years (Beer, 2009; Lash, 2007). In the past, many different systems and techniques have been developed to break down, order, and control multiplicities (e.g., bureaucratisation). Foucault (1980), for one, in his conceptualisation of the dispositif (p. 194), broadened the definition of power beyond rules or ordering to an understanding of relations of power entangled in infrastructures and mechanisms. The examination of the techniques of classification and infrastructures of arrangement has become increasingly relevant in the digital age.

Thinking of techniques of classification is conceptually useful for understanding the ways in which the distribution of the sensible is governed in social media platforms. Yeung (2017) uses the term "choice architectures" (p. 120) to describe how forms of algorithmic governance structure users' choices through what information, options, and suggestions are presented to them. Architecture is, and has historically been, used to describe modes of governance (Brighenti, 2010). For instance, Foucault's architectural model of power investigates the ways in which spaces are designed to make things visible in specific ways, and offers a useful parallel to analyse algorithms in social media architectural configurations. Foucault's configuration of power points to the multiple processes, measurements, calculations, and techniques, at play in organizing and arranging sociality—distributing bodies, lights, and gazes (Ziewitz, 2016). Analytically, this makes the sorting, ranking, and censuring, functions of algorithms specific techniques. Yet, this study also deviates from Foucauldian conceptualisations in some ways. While the regime of visibility inherent in the Panopticon is premised on the threat of constant visibility (Ziewitz, 2016), the regime of visibility online concerns the threat of (in)visibility —the unequal and uncertain distribution of visibility. Regardless, the management of visibility is undoubtedly an expression of power, not of one person over another in the Foucauldian sense, but rather of a technique of power making choices that shape the everyday experiences of billions of users. This realisation arguably diverts attention away from the merely functional understanding of algorithms, to an understanding of software as reality ordering. Moreover, through social media architecture, algorithmic governance also takes on an institutional character and becomes naturalised, constantly working in the background of our online space of interaction (Cristiano & Distretti, 2018; Flichy, 2013; König, 2019). This form of governance, first shaped by the processes of (in)visibilisation of users or information, orders reality online. This construction of reality shapes thought and action in our everyday lives. Online reality provides a picture of what can be conceived of and discussed, and thereon structures possibilities for thought. Algorithms invisibly enclose users, through personalisation, in very limited spheres of information which prevents them from accessing complex or contradictory information online. Consequently, it takes away from the processes of critical thought, hampering any reality beyond the one imposed online. Evidently, technology is not deterministic, and users have the potential to access complex information outside the social media environment. However, the ordering of reality online becomes problematic when its mechanisms are made invisible and the information is thereby presented as truth.[10]

The preceding considerations imply that algorithms, rather than being a simple set of code instructions, in actuality put into practice a specific vision of social order and truth. The effects of the techniques of classification and infrastructures of arrangement can only be detected through an overview of the system. In this light, reviewing the logics of social media platforms and engaging Foucault's theory lay ground for understanding the socio-political implications of inclusion and exclusion: (in)visibility. Rendering things visible, it is argued, is both a precondition and a tool of exercising power. If algorithms are altering our experience online, how do we ensure that different ontologies are present in how we experience social reality? For this, it is paramount to uncover the patterns of interaction of the different actors on social media that help sustain the regime of (in)visibility.

5 Chapter 2: Interaction, Co-Production, and Algorithmic Imaginaries

Algorithms do not function in isolation but are instead entangled in sociotechnical assemblages. "Assemblages" is a useful word for capturing the dynamics of the digital playing field. This is precisely because technology is so embedded within a multifaceted environment of social, cultural, and political interactions that it becomes a political concern, as it reflects particular ways of conceiving the world. Therefore, despite the previous chapter's emphasis on the material politics of algorithms, the materiality of platforms can never be detached from the domain of the social. As such, a sensitive socio-technical understanding is needed, to account for all the actors at play on internet platforms. A careful examination might provide some indication about how mechanisms of power come into being through platforms and their algorithms. At the intersection of algorithmic techniques, humans and experiences, this chapter is an approximation to the political and the bridge to the final chapter on resistance.

Social media platforms have become spaces where most

[10] Split Screens, a research project from The MarkUp, illuminates how Facebook's recommendation algorithm siloes information on the platform by collecting, analysing and putting side by side screenshots from the feeds of Democrat and Republican users in the United States.

of our sociality occurs online. Standardised modes of social relatedness across internet platforms include following other users, uploading everyday thoughts or experiences, reacting to other users' content (e.g., liking or retweeting), and tagging other users on pictures, to name a few. Social media platforms' penetration of the social fabric has gone in tandem with an increased significance of social media companies in the wider digital economy (Alaimo & Kallinikos, 2019). The shift towards datafication, the process of turning many aspects of daily life into data, has led to the transformation of social life into quantified online data (Cukier & Mayer-Schoenberger, 2020). In this way, internet platforms have become data-based organisations that exploit for profit the very sociality they host. In fact, the dynamics of data extraction have become so naturalised that they have become part of the inner workings of all social media platforms (Couldry & Mejias, 2019). More concretely, the platforms extract the "everydayness" from individuals' computer-mediated actions including Facebook "likes," text, photos, videos, location, communication patterns, purchases, and so on. Users' digital footprint is used to train algorithms with human data. These algorithms are dependent on the data to rank, order, and personalise content for engagement. More generally, the interaction that social media encourages, through various forms of user involvement, is collected to later be sold to advertising companies for targeting purposes (Alaimo & Kallinikos, 2017). Data extraction does not only make social relations profitable for corporations, but also allows for a new model of social governance through behavioural steering, essentially structuring the field of action of users (Couldry & Mejias, 2019). This, coupled with the imperative of data extraction for profit, creates systems where platforms aim to engage users constantly, with little regard for the content of the posts. This is particularly concerning because, as noted, algorithms create buzz around certain topics that attract high engagement, usually inflammatory posts. This type of content riles up people's reactions and therefore hampers in part the capacity of social media to become a space for constructive discussion on important topics. In fact, this became an important issue during the 2016 US presidential election's disinformation campaign. Russian trolls and bots designed posts with affective undertones, knowing that algorithms would make controversial content visible and that the content would heighten divisions between Republicans and Democrats ahead of the elections (Inkster, 2016).[11]

As touched upon in the previous chapter, algorithmic maps delineate patterns of user desirability, underlining questions about what behaviour is desirable online and what being a "good" user implies. Designers of social networks project specific meanings onto their platforms, "configuring" the imagined and idealised user (Velkova & Kaun, 2021). To remain visible, a "good" user must engage with algorithmic preferences, thereby participating on the platforms frequently and substantively. Investigating what kind of behaviours social media suggests and affords enables transcendence from the materiality of the platforms, so as to focus on the interactional outcomes. Bucher and Helmond (2017) propose the concept of "perceived affordances" (p. 5) to suggest that the designers of platforms should indicate how the user is meant to interact with the platform. This usually materialises in features available on the platform such as buttons, screens, or word limits on posts. Suggested forms of behaviour on social platforms give an indication of the operation of these spaces. Take for instance suggestions on Twitter: these include "trends for you" (Figure 7), "topics to follow" (Figure 8) and "who to follow" (Figure 9). Similarly, Facebook and Instagram suggest "people you may know" and allow you to choose from a number of emojis with which to engage—e.g., love, sad, angry, surprised, wow, haha, and care.

Figure 7: Trend suggestions on Twitter. This feature proposes topics of discussion that are popular in the user's geographical location.

Certain features such as the "like" button allow users to communicate with each other. For advertisers, these features facilitate preference data in order to tailor personalised ads and for platforms. For platforms on the other hand, the data allows them to measure engagement in order to personalise

[11] The disinformation campaign during the 2016 US Presidential elections is thought to have been carried out by Russian hackers with connections to the Russian intelligence Service (Inkster, 2016). During this campaign, Russian operations on social media supposedly took advantage of Facebook's power for the distribution of information and algorithmic logics so as to increase the scourge of "viral" hoaxes and create social divides ahead of the elections (Inkster, 2016).

Figure 8: Topic suggestions on Twitter. This feature allows users to select and follow trends they might be interested in. This also enables Twitter to tailor its suggestions to users.

Figure 9: User suggestions on Twitter. This feature lists friend suggestions based on the type of profiles a user follows as well as on their extended network (friends of friends).

each user's feed[12] (Bucher & Helmond, 2017). Hence, social media platforms are characterised by an architecture for engagement and their economic model, connecting user data to advertisers. More concretely, platforms enable and design the modes of sociality around their business model. In this way, social media platforms serve not only users but a plethora of actors, including platform owners, marketers, data analytics firms and partners (Alaimo & Kallinikos, 2017). This new form of capitalism is what Zuboff (2015) describes as 'surveillance capitalism', a system that aims to predict and modify human behaviour to produce revenue. When human experience is the input and output of these capitalist processes, it becomes problematic and raises questions about the power of platforms to exploit user interaction. The process of data extraction has been largely dehumanized when in reality, it is users' social and political engagement online that is being commoditized.

According to Foucault (1982), questions of power and exploitation cannot be studied in isolation to mechanisms of subjection, as they are intertwined in complex and circular ways. The extractive process that enables social media occurs in the absence of dialogue, typically running in the background of social interaction (Zuboff, 2015). In fact, the mere term extraction renounces any rationalisation of consent or reciprocity between social media companies and their users. Behavioural data online travels hidden paths of aggregation and translation to be decontextualised in code, even though it is immediately tied to actual people and human experience (Nissembaum, 2011). In this process, subjectivities are converted into objects that repurpose the subjective for commodification. Users' preferences and activity online is becoming increasingly visible to advertisers and platforms while the process of extraction remains largely invisible, only vaguely present in the terms and conditions of the platforms. In this way, visibility is a precondition for engaging in social media. In a Foucauldian vein, relationships of power on social networks are not necessarily a product of consent. Users are not directly agreeing to platforms shaping their possibilities of actions nor are they giving up full ownership of their data to be sold to private companies. In a similar manner, when cookie notifications pop up on a website, it is easier for users to accept them in passing instead of spending time setting their privacy preferences. However, accepting cookies is rarely an indication that users consent to their data being mined, rather it speaks more to the fact that cookies are often deceptive or confusing to the everyday user. Likewise, perhaps the voluntary participation of users with Internet platforms can be argued to indicate some form of prior consent. However, the relationship of power that plays out on social media is by no means the manifestation of a consensus (Foucault, 1982).[13]

Nonetheless, the digital subject breaks away from Foucauldian conceptualisations of the passive object of disciplinary power. Instead, the digital subject is thought of as active, making decisions and producing data, leaving digital

[12] The feed refers to the main screen on social media platforms which includes constantly updated posts from the people you follow as well as (product) advertisements.

[13] This also echoes Jacques Rancière's theorisations by Jacques Rancière. In his book Disagreement (1998), he states: "There is an order in society because some people command and others obey, but in order to obey an order at least two things are required: you must understand the order and you must understand that you must obey it. And to do that, you must already be the equal of the person who is ordering you" (p. 16). This is difficult in the case of social media where visibility asymmetries "invisibilise" the mechanisms of power embedded in platforms, thereby making consensus inconceivable.

traces that power social media (Krasmann, 2017). In fact, action is the very condition that enables subjection, which is ultimately the basis of the power relation. The visibility management of algorithms creates a 'fear of invisibility' in users which, paradoxically, seems to push users to continually engage with platforms' imperatives. According to Couldry and Turow (2014), platforms function as a public space. Precisely, visibility in social media allows users to participate in the construction of collective discourse and a political action, and this in turn, feeds the construction of users as social beings. Thus, invisibilising processes at the hands of algorithmic work entail the potential loss of recognition, and therefore the possibility of "mutual acknowledgement that orients us at least to recognise each other . . . as members of a common social and political space" (Couldry & Turow, 2014, p. 1,711). Therefore, the only way to escape from invisibility seems to be to participate in the capitalist logics these platforms promote. For example, users can buy followers or pay to create professional accounts on Facebook, Instagram, or Twitter, which ensures a greater diffusion of certain products, peoples, ideas, etc. More generally, users can choose to follow the platforms' suggestions, like following trends in "suggested for you." In this sense, the voluntary engagement of users with platforms actively generates data. Without this input,algorithms and ad personalisation cannot function. Algorithms have the ability to decide on what and who becomes visible on the feed. Moreover, algorithms categorise the individual, attaching to them their own identity and preferences, and impose a reality which users must recognise as true. In this manner, the power of algorithms brings subjects into being.

Yet, algorithmic power does not exist independent of subjects; they both interact and feed off each other to co-produce online reality. User-generated data must be produced and aggregated so that algorithms can make decisions on what appears on the feed (Ananny, 2016). For instance, the Facebook feature "On this day" (see Figure 10) employs an algorithm that continuously measures whether users interact with suggestions, as well as which memories the user would like to see on their feed (Bucher & Helmond, 2017). Similarly, the sorting, ranking, and censoring of content on Twitter is based on a "best Tweets first" logic built on users' own interaction with accounts they liked, tweets they retweeted, hashtags they used, etc. By clicking and liking, users feed the algorithms, which in turn, generate the information flows fed back to users. This generates a type of machine and human co-production where the algorithm and platform functions enable and restrain possibilities of visibility for users. At the same time, users, while engaging with social media, also influence algorithms, and in turn, what is going to appear on their feed. By engaging with platforms, users enable algorithmic work and contribute to their own subjection. Yet how can we formulate questions of power when reality online is co-produced?

According to Foucault (1982), power and freedom are not

Figure 10: 'On this day' feature on Facebook. The purpose of this tool is to remind a user of uploads or tags from years ago to encourage them to repost their memories.

mutually exclusive, but instead interact in complicated ways. For instance, although users are growing more aware of the extractive and privacy invasive logics of platforms, they still find pleasure in engaging online, sharing content, and gaining online popularity. This expresses the tension between visibility and invisibility that characterises the paradoxical relation of users with social media platforms. The relationship between power and freedom cannot be divorced since this would imply, according to Foucault (1982), "voluntary servitude" (p. 790),[14] a state that negates the possibility of power to constantly be renegotiated.

Nonetheless, underlying power asymmetries are still

[14]Foucault (1982) rejects the idea of voluntary servitude because, as he argues, "how could we seek to be slaves?" (p. 790)

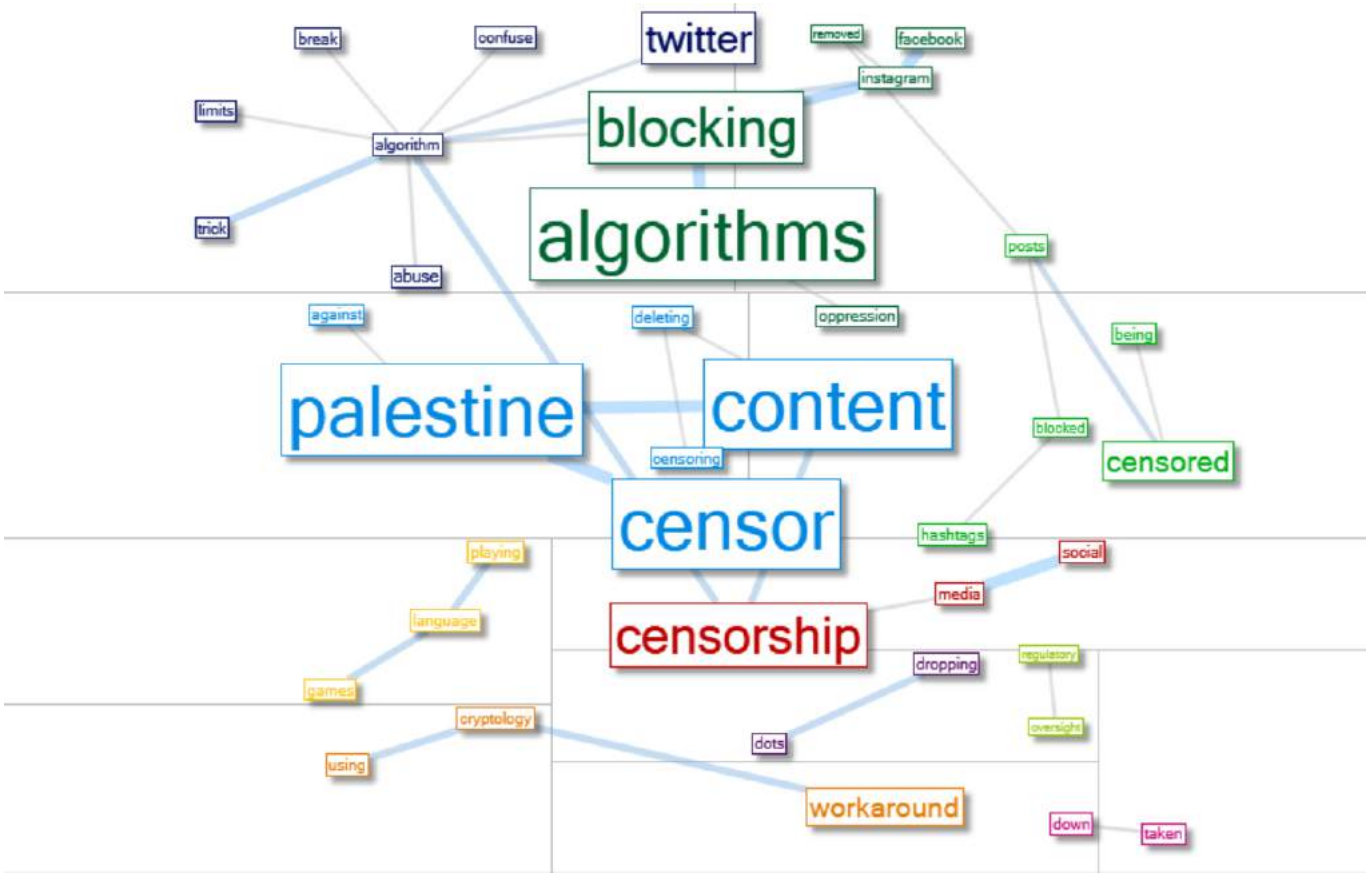

Figure 11: Semantic map surrounding the conversation on the algorithmic censorship of pro-Palestine content. The labels are sized according to the usage of words: the more times it is used, the bigger the label. Links represent connections between words. For instance, "posts" is being used with "censored"' but also with "Instagram." Created with NodeXL by searching "Palestine" and "algorithms" or "algorithm."

present on platforms. The hypervisibilisation of the user falls back on an infrastructure of invisible techniques that work to order reality online, raising important concerns about accountability and transparency. Thus, it is important to consider users not solely in terms of their interaction with the platform features, but also in terms of their experiences of how platforms work. The examination of this human-technology interaction can help to partly make visible the largely invisible workings of platforms. In particular, if we want to understand the power of algorithms, it is important to understand how users encounter and make sense of algorithms, and how these experiences, in turn, not only influence the expectations users have of platforms, but also help shape the algorithms themselves. This is what Bucher (2018) recognises as "algorithmic imaginaries" (p. 113); essentially, ways of thinking about what algorithms are and how they function. Rather than opening the "black box" by deciphering algorithmic code, the socio-technical assemblage in which they function allows us to understand their performance through the ways in which they are being articulated and experienced online. This raises interesting questions with regard to the power of algorithms and the ways in which they express this power, as well as how this may affect the environments in which they function. Essentially, these algorithmic imaginaries are derived from real-life human encounters with algorithms online. They are constructed through online conversations discussing how algorithms engage in different ways.

What is especially revealing is the moment when users turn their attention to algorithms that they perceive are influencing their online experience. Take for instance a recent example depicted in Figure 11.

This semantic map was created using a Twitter Search Network that included the words "algorithm" or "algorithms" and "Palestine." Essentially, the map represents the key words (size in relation to times-used) in a Twitter conversation around recent allegations that social media companies were blocking the visibility of posts related to Israeli violence on Palestinians, following the Israeli court verdict to forcibly expel Palestinian families from their homes in the Sheikh Jarrah neighbourhood (Alsaafin, 2021). The network demonstrates acute awareness of the workings of algorithms online as users attribute the takedown of pro-Palestinian content to algorithms and recognise their role in censoring, limiting and removing content as well as blocking users. The conversation highlights users' experience with Twitter, Instagram, and Facebook, and the actions they attribute to the

algorithmic work on these platforms. User sentiments such as "abuse," "oppression," and "regulatory oversight," express discontent with the algorithmic encounter. These negative sentiments are mostly centred on the concern that social media companies have a hidden political agenda (see Figure 12).

Figure 12: Tweet denouncing algorithms as biased in the context of the conflicts in Colombia and Israel/Palestine.

This discontent centres around the issue of visibility. The algorithm is tasked with distributing visibility of content and therefore, the attention it receives. We encounter reactions such as in Figure 13, where the user finds it "infuriating" that "the Instagram algorithm is hiding her stories," or another user who attributed getting "barely 65" views on her Instagram stories due to the platform's censorship logic (see Figure 14).

Figure 13: Tweet criticising Instagram's algorithm for hiding stories related to issues in the MENA region.

Figure 14: Tweet denouncing Instagram's censorship of Palestine-related content on stories.

This fear of invisibility or censorship reflects perceived power asymmetries and platform control that de-privilege users in relation to the platform. These vernacular accounts of how users come in contact with algorithms offer a preliminary indication of how users may approach these platforms. In fact, everyday encounters with algorithms can encourage users to develop tactics against algorithms. Some words like "workaround," "trick," "confuse," and "play" (as evidenced in Figure 12), already indicate a certain knowledge and predisposition to counter algorithmic work, drifting away from the "good" (data) subject imperative. Users increasingly recognise the role that they play in shaping these platforms in political, commercial, or playful ways (Velkova & Kaun, 2021).

The delineation of a "good" (data) subject through rules of engagement and possibilities of action assumes a passive agent who will do what is expected or suggested by the platform. Through the materiality of social media, we only know what the user is able to do. Conversely, through semantic network maps, we can inquire about how users use technology in practice and whether they will eventually resist or evade the configurations designated by algorithmic work.

As demonstrated, algorithms are not simply abstract computational processes; they also hold the power to enact material realities by shaping social life to various degrees. In our interaction with platforms, social media tells us something about ourselves and reflects the world in specific ways. Obviously, algorithms epitomise an unprecedented technologically-induced shift in our mode of world-making (Krasmann, 2017). What can be known and seen at a certain point in time on social media relies on a complex interaction of the material —techniques and architecture— and human users —their experience and participation. Therefore, to understand this kind of power, this chapter conceives of algorithms as functioning in assemblages of humans, political economies, and culture, in order to detangle each actor involved and understand the complex interrelations taking place on social networks. Through the analysis of human encounters with algorithms, the user is re-inserted in the formation of the complex socio-technical systems and acknowledges human-machine co-production, thereby refusing deterministic accounts of technology. The co-production of social reality online is where the question of the political setting, namely, what we see and what is perceived as truth. In this way, algorithmic power, which has taken a dystopian turn in a great part of the academic literature, can be largely demystified: it is no longer something invisible but rather something we can encounter in different ways; and no longer is it something that fully determines our social reality, but something we feed and can influence in different—yet largely unknown—ways.

6 Chapter 3: Resistance

The two previous chapters have shed light upon the consequences of algorithmically mediated visibility for the public sphere and digital citizenship. If political visibility and recognition is at play on social media platforms, the implications of algorithms go well beyond the ordering of visualities. Visibility management and dynamics of data extraction have pointed towards a decline of human agency online. The semantic map discussed in chapter two highlights user discontent with social media algorithms. Moreover, it denounces the infrastructure of invisible techniques that work to order reality online, raising important concerns about accountability and oversight. Yet, while the examination of user subjection and disempowerment through the themes of governance, control, and extraction, is central to existing research on power relations and internet platforms, instances of digital resistance remain underexplored. Therefore, this chapter sets out to extend the understanding of the user beyond its constant state of subjection and domination, and

explore everyday instances of users' algorithmic resistance.

This paper understands resistance as the examination of antagonisms against the effects of algorithmic power in relation to knowledge as well as the veil of secrecy that surrounds these techniques of power, hampering transparency, and thereby accountability. Resistance is algorithmic insofar as it refers to strategies where algorithms can be used as tools to enable this very resistance. Thus, algorithms are both mechanisms that enable subjection and mechanisms that enable possibilities of resistance. It is through the encounter of humans with algorithms that users may start devising strategies of resistance. In this way, it is the "algorithmic imaginary" that informs practices of resistance (Bucher, 2018, p. 113). In fact, the semantic map already points to some strategies being used in pro-Palestinian activism such as "language games," "using cryptology," and "dropping dots." These are the means through which ordinary users without technical knowledge navigate platforms, in a struggle to subvert the technical mechanisms that are structuring their everyday lives online.

Since algorithms assume a dominant role in the mediation of online reality, this study seeks to uncover algorithmic expressions of power through the architectural and interactional aspects of the regime of visibility conceptualised in Chapters One and Two. Finally, as proposed by Foucault (1982), this chapter examines power relations of user and algorithm through practices of resistance. Foucault (1982), in his shift from power to the subject in his later work, suggests this "new economy of power" as more empirical, bridging theory and practice and thereby directly related to the current situation.

In the same way co-production acknowledges an active digital subject, resistance presupposes an active subject. Without the power to act, the user cannot hold the potentiality to intervene against repressive mechanisms of power (Foucault, 1982). However, conceptualisations of resistance are not limited to one kind of subject, rather this study recognises multiple and multifaceted modes of subjectivity and agency present in the act of resistance. As Foucault (1982) contends, inherent in the condition of power relations, there exists modes of insubordination in its "essential obstinacy on the part of the principles of freedom" (p. 794) as well as "modes of escape or flight" (p. 794), both of which hold the potentiality of struggle. In this way, it is not about limiting our understanding of resistance strategies as a confrontation between domination and freedom but rather, about grasping its complicated interplay. For instance, while some might find resistance in using algorithmic mechanisms to their advantage, other non-complicit forms of resistance might simply involve opting-out of the platforms altogether. And others might not even be able to afford revolting against these platforms altogether as they struggle to remain out of sight of surveillance systems. In fact, revolting against algorithmic systems is a privilege in itself. For instance, many activists in authoritarian states cannot afford to resist as they fear reprimand from the government or heightened surveillance (Kaun & Trere, 2020).

By reason of these intricacies, we must recognise the plethora of practices through which people act with, on, and against social media platforms. Instead of only listing possible resistance tactics, this study aims to offer a broad classification of types of resistance as a way of grouping. These include instances of both active resistance—semantic resistance, hashtag takeovers, redeploying and tricking the algorithm—and passive or non-compliant resistance—disconnection and obfuscation. Through an empirical engagement with resistance case studies, power operations can be highlighted, and the socio-political implication of algorithms further determined.

Active resistance

Active resistance is a complicit form of resistance. It does not deny the power of algorithms but rather operates within their frameworks, using them and other infrastructural features for different ends or to users' advantage. This way of making use of the affordances of digital platforms allows digital subjects to construct spaces of resistance online, while also affecting other elements of the digital environment. Active resistance strategies are in a perpetual process of reimagination , as the terms and conditions of platforms and algorithms are constantly tweaked by their designers.

Semiotic Resistance: Semiotic resistance broadly refers to a common linguistic practice of using non-standard spellings in posts to evade algorithmic censure or imposed invisibility when posting about topics that do not respect the often-vague or invisible community guidelines of internet platforms. Content moderation algorithms claim to mitigate false or harmful content but largely seem to police the language of posts uploaded for political purposes. As such, non-standard spellings are useful in reducing the visibility of the post vis-à-vis algorithmic censorship—and sometimes human moderators—while increasing the visibility of the post in those networks of users where the alternative spelling has accrued a specific meaning.

Often referred to as neural text attacks, these strategies attempt to modify text in ways that are imperceptible to humans but will cause algorithms to detect the text as human-written (Wolff & Wolff, 2020). In other words, unlike humans, algorithms that are programmed to censor textual content that violates community guidelines do not "understand" that visually, similar letters can be interchangeable to human readers. While this does not change the message for fellow users, it will confuse many machine learning algorithms.

There are two types of neural text attacks. The first are non-human attacks that change text in ways a human

normally would not (Wolff & Wolff, 2020). This is done through replacing certain letters with homoglyphs or symbols. The most common example in social media is the phonetic matching of English letters with Cyrillic letters (see Figures 15 and 16).

Figure 15: Tweet explaining how to use Cyrillic alternatives to avoid censorship of certain words.

Figure 16: Tweet using Cyrillic letters to avoid censorship of the word 'fuck'.

Another recent example is a tactic referred to as dotless Arabic script, an ancient method of writing Arabic which is being used to combat the censorship of content regarding the situation in Palestine. Arab-speaking users are opting to write posts in a form of Arabic without dots to navigate the censorship of algorithms and spread pro-Palestinian content (see Figure 16). To facilitate this task, hundreds of dotless keywords and apps have been developed and are openly available on the internet.[15] Since the escalation of violence between Israel and Palestine in May 2021, a report by the Arab Centre for the Advancement of Social Media, known as 7amleh, has showed a significant increase in the censorship of pro-Palestinian content and accounts trying to raise awareness about the situation. The report concluded 50% of these reported incidents occurred on Instagram, 35% on Facebook, 11% on Twitter, and 1% on Tik Tok ('The attacks on Palestinian digital rights', 2021). Consequently, users have had to learn to work around the mechanisms of algorithms that flag content as inappropriate and remove it. In this way, creative uses of cryptology with easy-to-use tools have been developed.

These alternative uses of language are acquiring meaning and functioning in a sort of language of rebellion around the common cause of Palestine (see Figure 17).

The second type of neural attack is human-like, meaning that it involves changing text in a way that humans do, for instance, misspelling words or changing punctuation or grammar (Wolff & Wolff, 2020). This has been another common tactic against the censorship of pro-Palestine concepts, with alternative spellings such as "Pal-estine," "Filistine,"

[15] See https://www.dotless.app/ or https://m.apkpure.com/dot-less-arabic-keyboard/com.dotless.arabic.keyboard

Figure 17: Facebook post which uses dotless Arabic.

Figure 18: 'F*ck your algorithm' dotless Arabic illustration

"G@za," "Israhell," or "¡srae!," non-standard spellings that are thought to trick the algorithm (see Figure 18).

Figure 19: Tweet containing human-like neural attack with alternative spellings "Israhell" and "G@za."

This practice is by no means new: it was widely used during the Black Lives Matter (BLM) movement in 2019 with spellings such as "Wypipo," "yt," and "Whyte" to resist an algorithmic bias that flagged "whites" or "white people" as "protected categories" (Angwin, 2017). As explained in Chapter Two, algorithms are not neutral, rather they are reflective of social norms and biases. The flagging of posts

identifying "white people" as hate speech contributes to an ongoing narrative where algorithms, which are supposed to protect minority or marginalised groups by removing hate speech, end up silencing them. Interestingly, when interrogated about the mass takedowns during the BLM movement and during the recent Israel/Palestine crisis, social media companies have attributed the controversial censorship to a "glitch in the system" (Gebeily, 2021), in an attempt, it seems, to reassert an unbiased and mechanic image of algorithms. The deliberate ignorance of the socio-political power of algorithms online, i.e., framing algorithms as objective or mechanized allows Tech Giants to evade accountability to serve their economic interests while also avoiding external regulation.

Figure 20: Facebook post explaining alternative spelling for "white."

Figure 21: Tweet containing the alternative "yt" spelling.

As such, semiotic resistance deploys "figurative language, semiotics, indirectness and wordplay as a means of conveying multiple layers of meaning" (Florini, 2014, p. 1). These strategies craft collective understandings of the meanings of alternative spellings and often create possibilities for vulnerable or marginalised users to communicate and for their causes to be "seen."

Hashtag Takeover: Hashtag takeover strategies refer to the redeployment of a hashtag and the redefinition of its meaning by users, in ways that were not initially intended. This is usually done by liberal activists in order to reinvent hashtags they consider hateful or discriminatory. In this way, activists increase the visibility of political causes without challenging the algorithmic logic as such (Treré, 2017). This is a common practice on Tiktok, Twitter, and Instagram, with activists defusing racist hashtags or embedding radical messages, and working around the algorithm to promote alternative messages. Some notorious examples are #WhiteLivesMatter,[16] #BlueLivesMatter,[17] and #WhiteOutWednesday[18] hashtag takeover where traditional activists and K-Pop fans forged alliances to drown out racist or offensive posts. K-Pop fans would flood these tags with fancams, videos and memes of K-Pop music groups (see Figure 21).[19] In this way, these hashtags lost their intended meaning. Even the posts that were still using the tags to promote racist content would get obfuscated by K-Pop content.

Figure 22: Picture of K-Pop artist with #WhiteLivesMatter hashtag.

Another recent example is the takeover of the hashtag #ProudBoys (see Figure 22). Proud Boys is a far-right group that took part in the coordination of the Capitol attack, and that was labelled a terrorist group by the Canadian government (Feuer, 2021). The hashtag, which was used to incite violence around the Capitol Hill incident, has now been redeployed by the LGBTQIA+ community and is being used to post queer content.

This strategy is also currently being used by pro-Palestinian activists who have taken over the hashtag #VisitIsrael, a hashtag originally intended as part of a campaign to attract tourism to Israel (see Figure 23).

Evidently, this type of resistance is only temporary. Activists need to scout for and engage with countermovement's hashtags that are constantly being reinvented. However, this type of collective and organised algorithmic resistance demonstrates means of mobilising in tune with algorithms toward ends that were not originally intended (Velkova & Kaun, 2021). Ultimately, for liberal activists, hashtag takeovers function as a way to divert political attention away from the hateful content that algorithms are not programmed to censor. Instead, groups of users try to decrease the visibility of this content and to amplify political causes deemed more deserving of attention.

[16] White supremacist hashtag that emerge in reponse to the Black Lives Matter Movement in 2015. It was first used by he Texas-based white supremacist group Aryan Renaissance Society and later adopted by the Ku Klux Klan in 2016.

[17] Hashtag that accompanied the Black Lives Matter Movement that arose in 2015 to protest against police brutality against African-Americans in the United States.

[18] Hashtag that arose to undermine the #BlackOutTuesday hashtag which was intended to honour the loss of African-American lives to police brutality.

[19] Fancams are usually videos filmed by a member of the live audience during a music performance.

Figure 23: Tweet commenting on the #ProudBoys LGBTQIA+ takeover.

Figure 24: Tweet denouncing Israeli violence on Palestinians using the #VisitIsrael hashtag.

Playing Algorithms: As they currently are designed, content moderation systems, powered by flagging algorithms, remove content at a massive scale, which can be extremely frustrating to vulnerable communities or activists trying to raise their voices about an issue. Several practices have been developed that are thought to evade or confuse algorithms and, thereby, avoid censorship. The most used approach to remove image-based content is called the "checksum": an algorithm that computes whether two images are identical in a matter of seconds. Hence, checksum is highly instrumental for mass takedowns around a certain issue. The checksum mechanism can be misled by either introducing noise to the image (visually distorting the image with a grain-like effect) to be posted or by cropping the image to a slightly different size than the original (as evidenced in Figure 24). Although more technical, these strategies are accessible to ordinary users. In fact, there are free editing softwares online that can automatically do this[20].

Video censorship—enabled by speech-to-text algorithms—is more difficult to work around because audio editing software is not as widely available. However, a weakness in these algorithms is that they struggle to identify speech when there are multiple speakers talking at once. Therefore, users can inhibit algorithmic performance by playing a podcast or some other audio recording in the background while speaking. While a human can distinguish a person's words from background noise, algorithms struggle to do so. Moreover, these algorithms require a lot of data on a language to transcribe speech to text, so speaking a dialect or an accent will also inhibit their performance. For instance, different dialects of Arabic should be very difficult to detect as there exists little data on all different forms of Arabic.

[20]The images in figure one were edited using free software found on: `https://pinetools.com/add-noise-image`

Figure 25: Image before and after adding noise. *The image is still visible to the human eye, but confuses algorithms.*

Avoiding the censorship of visual media is essential to online activism, especially in movements that involve violence, police brutality, and other human rights abuses. As Neumayer and Rossi (2018) argue, images and videos "are key to rendering particular events and perspectives on political contestation visible or invisible" (p. 4,295). Indeed, they create visual narratives—an expression of dissent, a type of eye witnessing, or an alternative to mainstream outlets' narratives of the events. Images can also serve as reference points to indicate where riots and other forms of violence are taking place. This was the case during the Black Lives Matter movement: users organised demonstrations through Facebook, Instagram, and Twitter, and gave information about the protests through Instagram stories. Moreover, visuals can also be utilised for the purpose of putting up photos of missing, arrested, or killed people. This is especially important on social media, where people all over the world remotely join the conversation on

what is happening in a physical space.

During the 2019 protests in Chile[21], there were mass take-downs of police brutality content documented by the Fundación Datos Protegidos[22], mostly coming from Instagram algorithms. The censorship of this content implied the loss of material for future visualisation, testimony, and memory of the incidents that occurred. In order to counter censorship and possibly the erasure of visual history, a Twitter account by the name of @ArchivandoChile compiled all images and videos of protests to an archive in a communal Telegram channel[23], thereby creating a public registry of the events and a collective visual memory (see Figure 25).

Pro-Palestinian activists are instead using a "sandwiching posts" tactic. Users started to complain that their Instagram stories with Palestinian content were getting drastically less views than their normal posts. To counteract this, activists have learnt to post random content in between Palestinian posts on their stories. For instance, posting pictures of food, random objects, or makeup videos to confuse the algorithm into thinking that they are posting non-political content.

These two contextual examples demonstrate how contemporary digital activism increasingly means understanding how algorithmic systems work. Users are developing counteracting techniques to "play" the algorithm, thereby meeting their socio-political needs in the name of social justice or political change. Virtually, technology is being envisioned as an extension of the site of struggle.

Passive resistance

Passive resistance is a form of non-compliant resistance. This means that rather than conforming to the power of algorithms, users refuse it entirely. Little attention has been drawn to the ways in which users are purposefully choosing to (partially) disconnect from or obfuscate platforms as an act of resistance, and are crafting alternatives to constant connectivity. As Syvertsen (2017) describes, "negative actions and attitudes towards media describe refusal to accept the way media operates and evolves" (p. 9) therefore understanding the act of disconnection as political in nature.

Disconnection: Growing criticism of content moderation

Figure 26: Tweet in Spanish that reads: 'Creating a free memory without censorship. Mention @ArchivandoChile when you share an image or video, and it will be kept for public record and without bias of what is happening in Chile today'

[21] The 2019 Chilean protests -or the "Estadillo Social" as it is known in Chile-, were a series of massive demonstrations and riots against social inequality and demanding the resignation of President Piñera.

[22] Fundación Datos Protegidos -Protected Data Foundation in English- is a Chilean NGO created in 2015 with the objective of promoting data protection and privacy. Report: https://datosprotegidos.org/descarga-informe-libertad-de-expresion-en-el-contexto-de-las-protestas-y-movilizaciones-sociales-en-chile-durante-el-estallido-social/

[23] Telegram is a free and open-source instant messaging app. The platform provides end-to-end encrypted messaging, voice, and video calls. It is often used as a more secure alternative to WhatsApp. In fact, Telegram is well known as a platform where people contact drug dealers to buy drugs.

Figure 27: Tweet explaining the sandwiching technique.

on social media and the surveillance capitalist business model of digital companies has made users increasingly aware of the price that comes with constant connectivity (Zuboff, 2021). In the midst of the current Covid-19 pandemic, the lines between the offline and online state have been blurred, leading instead to a state of hyperconnectivity. More and more, users are opting to reject social media's connective affordances and to withdraw partially or wholly from the sites (Kaun & Treré, 2020). These practices can be temporal as they are constantly being renegotiated against platform developments -conditions of use and the political context.

For example, when WhatsApp announced its change of terms back in January, it prompted a backlash from users. The new terms involved WhatsApp sharing data with Facebook, who purchased the instant messaging platform in 2014. Though messaging remained encrypted, meta-data and contacts are being shared with Facebook. The lack of clarity around what kind of data was being extracted and shared by WhatsApp infuriated many users. This incident led to some users moving to other platforms such as Signal[24] and Telegram, with Signal being downloaded 246,000 times worldwide in the week before WhatsApp announced the change, and 8.8 million times in the week after (O'Flaherty, 2021).

Beyond examples of full disconnection, there have also been recorded instances of partial disengagement. Goyanes and Skoric (2021) researched the social media engagement of Catalans during the height of the independence movement. They found out that Catalan users decided not to engage actively in online debates about the Catalan referendum but rather to be present on social media as mere spectators to keep informed (Goyanes & Skoric, 2021). Many users choose to avoid open and public engagement out of fear of reprisal in their social and professional life (Goyanes & Skoric, 2021). Instead of heightening their visibility online, they choose to remain invisible, many opting to talk about the topic in-person instead (Goyanes & Skoric, 2021).

Yet, disconnection is not always a matter of choice. Countries with poor infrastructure or low internet signals are simply not able to connect (Treré et al., 2020). In other situations, disconnection can also be forced when authoritarian states impose strict state surveillance and control over the internet and users fear repercussions (Treré et al., 2020). Finally, in a world of constant connectivity, sometimes opting out is not an option. Disconnection can prompt a loss of contact with friends and family and can even be detrimental to finding job prospects. This is especially true during the current Covid-19 pandemic where most of our daily activities take place online. Many people are dependent on tools like Slack or Zoom to carry out their daily work, study, or social activities.

Obfuscation: Another mode of non-compliant resistance aims to obfuscate algorithmic systems of subjection. Most practiced is the method of data obfuscation (Ettlingler, 2018). In an era where our attention is systematically exploited through increasingly algorithmic techniques, this issue acquires an even stronger political relevance. Reports about privacy breaches and data extraction on social network sites have led savvy users to use a Firefox extension called FaceCloak as an obfuscation strategy[25]. FaceCloak is an architecture that protects users' privacy on a social networking site by shielding their personal information from the site and from other users that were not given explicit authorisation (Luo et al., 2009). To do this, it generates fake information Facebook requires for your profile and stores sensitive information in encrypted form on a separate server (Luo et al., 2009). FaceCloak also allows users to be selective about what information they want to protect, tailoring the software to serve users' own degrees of privacy (Luo et al., 2009).

Users that engage in these strategies demonstrate acute knowledge about the functioning of power in social media and purposefully choose to not work within its frameworks. Although more research is needed on strategies of passive resistance, it is possible that these practices will become increasingly common. Algorithms are not static, but in constant development. Consequently, they can be tweaked to adapt to active resistance strategies in order to neutralise or co-opt resistance efforts. It seems that social media companies are starting to find ways to counteract user antagonism. For example, users are now reporting that the "sandwiching technique" is no longer effective (Figure 26) for avoiding censorship on pro-Palestine content. Moreover, hashtags in solidarity with Palestine are being updated because users report that the algorithm is censoring the original ones: #Gaza_Under_Attack and #FreePalestine. Against this backdrop, there is a possibility that users will instead choose to disconnect or obfuscate from platforms as more productive and lasting forms of resistance instead of continuously attempting to battle the ever-evolving algorithms.

Figure 28: User reporting that the sandwiching technique is no longer effective.

In conclusion, active and passive practices demonstrate multiple, and multifaceted means of resistance. While users

[24]Signal is another free and open-source instant messaging app. In most aspects it is similar to Telegram, providing end-to-end encryption on the chat.

[25]Software can be installed online at `https://crysp.uwaterloo.ca/software/facecloak/`

might interact within algorithmic frameworks in order to bring heightened visibility to their content, other users might opt or be forced to refuse connectivity altogether, thereby remaining invisible. The persistent re-imagination of strategies calls for critical reflection on the autonomy of users online and their ability to adapt to the technological present. This does not only pose challenges to the dominant power of social media and its regime of visibility, but also allows for imagining new and better possibilities for these digital spaces. This thesis does not suggest that the responsibility should fall on the user alone to try and counteract the power asymmetries. A genuine effort should be made by companies to redefine the parameters of what algorithmic success is and whether the functioning of these is serving society at large. In order to reconcile opposing interests, state, market, and civil society actors need to renegotiate the values embedded in the design of platforms. Moreover, policies and regulations should hold digital platforms accountable and continue to be developed. However, in the functioning of these spaces, as demonstrated, the user is of great importance in holding the politics of algorithms accountable. The previous two chapters have established the dynamics of subjection and control embedded in platforms. The interplay of algorithms and strategies of resistance offers new insights into these dynamics, charting productive approaches for the management of algorithmic power in everyday life.

7 Conclusion

This study provided a new conceptualisation of a social media regime of visibility so as to explore its socio-political implications, in theoretical and empirical ways. Through a novel analysis of network maps and strategies of algorithmic resistance online, this research contributes to uncovering the power of algorithms, inscribed as "black box" technologies in the field of critical media studies (Pasquale, 2015).

This study has identified how algorithmically-managed visibility, enabled by the architectural and interactional aspects of social media platforms, is giving rise to symbolic forms of control, subjection, and power. The first two chapters on the architectural and interactional aspects of the regime established the claims for algorithmic power. Through an analysis of social media network maps, Chapter One uncovered patterns of (in)visibility that delineated the dynamics of inclusion and exclusion embedded in platforms. Ultimately, it is argued, the architecture of internet platforms create systems of governance that shape social reality online. Expanding on this, Chapter Two touches upon political economies and data extraction, further delving into subjection and what it means to be a "good" (data) subject on internet platforms. Through an analysis of a semantic map, it contributes to the demystification of algorithmic work, by arguing that social reality is in fact partly co-produced by humans and technology. Finally, informed by the theoretical backbone of the previous chapters, Chapter Three contextualises the research by examining instances of active and passive algorithmic resistance, thereby revealing users' expanding awareness of algorithmic power and growing demands for accountability.

Certainly, the struggle for online visibility has gained contours under this new conceptualisation of a regime. Through the concept of (in)visibility, this research was able to devise the implicit tensions between user subjection —extraction and behavioural steering—and user autonomy—coproduction and resistance. Moreover, (in)visibility proved to be a useful concept in bridging user experience on platforms with theoretical underpinnings and uncovering socio-political implications. Notably, it uncovered the ways in which social reality online can have real life implications, especially regarding knowledge. Although algorithms function as invisible techniques, this paper delineates the ways in which knowledge circulates, mediated by technologies in sorting, ranking, and censoring ways. Finally, through the themes of technique and data extraction, this paper reinforced the importance of demanding more transparency and accountability from internet platforms, as their power continues to function largely unscrutinised.

Essentially, what the conceptualisation of a new regime of visibility online has shown is that in the end these spaces which initially offered incredible possibilities to access information and for political action (and still hold the potential for it), are in fact machines of reproduction of the inegalitarian social order. Even more problematic is the fact that they are in the hands of a very limited number of actors -Tech Giants[26]- who abuse their dominant position to make extraordinary profits which are not reinvested to rebalance the existing inequalities in society, but on the contrary, reinforce them.

To counteract this grim picture, more research is needed to understand the functioning of algorithms and strategies of ordinary resistance. In this way, users can learn to navigate and negotiate both the space of visibility and invisibility without succumbing to the exploitative logics of platforms. Meanwhile, a real effort should be made from governments and Tech Giants to redefine the parameters of what algorithmic success is and whether the functioning of these techniques make sense for society at large.

References

The attacks on Palestinian digital rights. (2021, May 21). *7amleh: Arab Centre for the Advancement of Social Media.* https://7amleh.org/2021/05/21/7amleh-issues-report -documenting-the-attacks-on-palestinian-digital-rights.

Alaimo, C., & Kallinikos, J. (2017). Computing the

[26]These include among others all major social media platforms: Instagram, Facebook, Twitter, etc.

everyday: Social media as data platforms. *The Information Society, 33*(4), 175–191. https://doi.org/10.1080/01972243.2017.1318327

Alaimo, C., & Kallinikos, J. (2019). Social media and the infrastructuring of sociality. In M. Kornberger, G. C. Bowker, J. Elyachar, A. Mennicken, P. Miller, J. R. Nucho, & N. Pollock (Eds.), Research in the sociology of organizations (pp. 289–306). *Emerald Publishing Limited.* https://doi.org/10.1108/S0733-558X20190000062018

Albu, O. B., & Hansen, H. K. (2021). Three sides of the same coin: Datafied transparency, biometric surveillance, and algorithmic governmentalities. *Critical Analysis of Law, 8*(1), 9–24.

Alsaafin, L. (2021, May 1). What is happening in occupied East Jerusalem's Sheikh Jarrah? *Aljazeera*. https://www.aljazeera.com/news/2021/5/1/what-is-happening-in-occupied-east-jerusalems-sheikh-jarrah

Ananny, M. (2016). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology, & Human Values, 41*(1), 93–117.

Argwin, J. (2017, June 28). Facebook's secret censorship rules protect White men from hate speech but not Black children. *ProPublica*. https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society, 11*(6), 985–1002. https://doi.org/10.1177/1461444809336551

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society, 20*(1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147

Berger, P. L., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge,*. Penguin P.

Brighenti, A. M. (2017). The visible: Element of the social. *Frontiers in Sociology, 2,* 17. https://doi.org/10.3389/fsoc.2017.00017

Brunton, F., & Nissenbaum, H. F. (2016). *Obfuscation: A user's guide for privacy and protest.* MIT Press.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society, 14*(7), 1164–1180. https://doi.org/10.1177/1461444812440159

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society, 20*(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086

Bucher, T. (2018). *If ... then: Algorithmic power and politics.* Oxford University Press.

Bucher, T., & Helmond, A. (2018). *The affordances of social media platforms.* Sage Publications. https://dare.uva.nl/search?identifier=149a9089-49a4-454c-b935-a6ea7f2d8986

Cardon, D. (2018). The power of algorithms. *Pouvoirs, No 164*(1), 63–73.

Couldry, N., & Mejias, U. A. (2019). Making data colonialism liveable: How might data's social order be regulated? *Internet Policy Review, 8*(2). https://doi.org/10.14763/2019.2.1411

Couldry, N., & Turow, J. (2014). Advertising, big data and the clearance of the public realm: Marketers' New Approaches to the Content Subsidy. *International Journal of Communication, 8*, 1-17.

Crawford, K. (2016). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values, 41*(1), 77–92. https://doi.org/10.1177/0162243915589635

Cristiano, F., & Distretti, E. (2020). Toward an aesthetics by algorithms: Palestinian cyber and digital spaces at the threshold of (in)visibility. In *The aesthetics and politics of the online self* (pp. 1–34). Palgrave Macmillan.

Cukier, K. N., & Mayer-Schoenberger, V. (2020, May 1). The rise of big data. *Foreign Affairs.* https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data

DeVito, M. A., Gergle, D., & Birnholtz, J. (2017). 'Algorithms ruin everything': #RIPTwitter, folk theories, and resistance to algorithmic change in social media. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems,* 3163–3174. https://doi.org/10.1145/3025453.3025659

Dijck, J. van, & Poell, T. (2013). Understanding social media logic. *Media and Communication, 1*(1), 2–14. https://doi.org/10.17645/mac.v1i1.70

Ellerbrok, A. (2010). Empowerment: Analyzing technologies of multiple variable visibility. *Surveillance & Society, 8*(2), 200–220. https://doi.org/10.24908/ss.v8i2.3486

Ellul, J. (2006). *The technological society*. Vintage.

Ettlinger, N. (2018). Algorithmic affordances for productive resistance. *Big Data & Society, 5*(1), 205395171877139. https://doi.org/10.1177/2053951718771399

Feuer, A. (2021, February 5). Did the proud boys help coordinate the Capitol riot? Yes, U.S. suggests. *The*

New York Times. https://www.nytimes.com/2021/02/05/nyregion/proud-boys-capitol-riot-conspiracy.html

Flichy, P. (2013). Rendre visible l'information. *Reseaux, n° 178-179*(2), 55–89.

Florini, S. (2014). Tweets, tweeps, and signifyin': Communication and cultural performance on "Black Twitter." *Television & New Media, 15*(3), 223–237. https://doi.org/10.1177/1527476413480247

Foucault, M. (1982). The subject and power. *Critical Inquiry, 8*(4), 777–795.

Foucault, M. (1989). *De la gouvernementalité: Leçons d'introduction aux cours des années 1978 et 1979*. Ed. du Seuil.

Foucault, M. (1995). *Discipline and punish: The birth of the prison /*. Vintage Books.

Foucault, M., & Gordon, C. (1980). *Power/knowledge: Selected interviews and other writings, 1972-1977*. Pantheon Books.

Foucault, M., & Senellart, M. (2011). *The birth of biopolitics lectures at the college de France, 1978-1979*. Palgrave Macmillan.

Gebeily, M. (2021, May 10). Instagram, Twitter blame glitches for deleting Palestinian posts. *Reuters*. https://www.reuters.com/article/israel-palestinians-socialmedia-idUSL8N2MU624

Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies* (pp. 167–194). The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Goyanes, M., & Skoric, M. (2021). Citizen (dis)engagement on social media: How the Catalan referendum crisis fostered a teflonic social media behaviour. *Mediterranean Politics*, 1–22. https://doi.org/10.1080/13629395.2021.1904349

Hansen, H., K., & Flyverbom, M. (2015). The politics of transparency and the calibration of knowledge in the digital age. *Organization, 22*(6), 872–889.

Inkster, N. (2016). Information warfare and the US presidential election. *Survival, 58*(5), 23–32. https://doi.org/10.1080/00396338.2016.1231527

Just, N., & Latzer, M. (2017). Governance by algorithms: Reality construction by algorithmic selection on the Internet. *Media, Culture & Society, 39*(2), 238–258. https://doi.org/10.1177/0163443716643157

Kantrowitz, A. (2016, February 6). Twitter to introduce algorithmic timeline as soon as next week. *BuzzFeed*. https://www.buzzfeednews.com/article/alexkantrowitz/twitter-to-introduce-algorithmic-timeline-as-soon-as-next-we

Kaun, A., & Treré, E. (2020). Repression, resistance and lifestyle: Charting (dis)connection and activism in times of accelerated capitalism. *Social Movement Studies, 19*(5–6), 697–715. https://doi.org/10.1080/14742837.2018.1555752

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society, 20*(1), 14–29. ttps://doi.org/10.1080/1369118X.2016.1154087

König, P. D. (2020). Dissecting the algorithmic Leviathan: On the socio-political anatomy of algorithmic governance. *Philosophy & Technology, 33*(3), 467–485. https://doi.org/10.1007/s13347-019-00363-w

Krasmann, S. (2017). Imagining Foucault. On the digital subject and "visual citizenship." *Foucault Studies*, 10–26. https://doi.org/10.22439/fs.v0i0.5339

Lash, S. (2007). Power after hegemony. *Theory, Culture and Society, 24*(3), 55–78. https://doi.org/10.1177/0263276407075956

Luo, W., Xie, Q., & Hengartner, U. (2009). FaceCloak: An architecture for user privacy on social networking sites. *Proc. of 2009 IEEE International Conference on Privacy, Security, Risk and Trust (PASSAT-09)*, 26–33.

Musiani, F. (2013). Governance by algorithms. *Internet Policy Review, 2*(3). https://doi.org/10.14763/2013.3.188

Neumayer, C., & Rossi, L. (2018). Images of protest in social media: Struggle over visibility and visual narratives. *New Media & Society, 20*(11), 4293–4310. https://doi.org/10.1177/1461444818770602

Neumayer, C., Rossi, L., & Struthers, D. M. (2021). Invisible data: A framework for understanding visibility processes in social media data. *Social Media + Society, 7*(1), 205630512098447. https://doi.org/10.1177/2056305120984472

Nissenbaum, H. (2011). A contextual approach to privacy online (SSRN Scholarly Paper ID 2567042). *Social Science Research Network*. https://papers.ssrn.com/abstract=2567042

O'Flaherty, K. (2021, January 24). Is it time to leave WhatsApp – and is Signal the answer? *The Guardian*. www.theguardian.com/technology/2021/jan/24/is-it-time-to-leave-whatsapp-and-is-signal-the-answer

Pasquale, F. A. (2015). The black box society: *The secret algorithms that control money and information*. Digi-

talCommons@UM Carey Law. http://www.hup.harvard.edu/catalog.php?isbn=9780674368279

Poell, T., Nieborg, D., & Dijck, J. van. (2019). Platformisation. *Internet Policy Review, 8*(4). https://policyreview.info/concepts/platformisation

Rancière, J. (2013). *The politics of aesthetics: The distribution of the sensible.* Bloomsbury.

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results. *Convergence, 24*(1), 50–68. https://doi.org/10.1177/1354856517736982

Smith, H., & Burrows, R. (2021). Software, sovereignty and the post-neoliberal politics of exit. *Theory, Culture & Society,* 026327642199943. https://doi.org/10.1177/0263276421999439

Syvertsen, T. (2017). *Media resistance: Protest, dislike, abstention.* Palgrave Macmillan.

Thrift, N. (2004). Remembering the technological unconscious by foregrounding knowledges of position. *Environment and Planning. D, Society & Space., 22,* 175–190.

Treré, E., Natale, S., Keightley, E., & Punathambekar, A. (2020). The limits and boundaries of digital disconnection. *Media, Culture & Society, 42*(4), 605–609. https://doi.org/10.1177/0163443720922054

Treté, E. (2018). From digital activism to algorithmic resistance. *In The Routledge Companion To Media And Activism* (pp. 367–375). Routledge. https://doi.org/10.4324/9781315475059-39

Velkova, J., & Kaun, A. (2021). Algorithmic resistance: Media practices and the politics of repair. *Information, Communication & Society, 24*(4), 523–540. https://doi.org/10.1080/1369118X.2019.1657162

Vickers, B., & McDowell, K. (Eds.). (2021). *Atlas of AI.* Ignota Books.

Weiskopf, R. (2021). Dis/organising visibilities: Governmentalisation and counter-transparency. *Organization,* 135050842199575. https://doi.org/10.1177/1350508421995751

Wolff, M., & Wolff, S. (2020). Attacking neural text detectors. *ArXiv:2002.11768 [Cs].* http://arxiv.org/abs/2002.11768

Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society, 20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Ziewitz, M. (2016). Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values, 41*(1), 3–16. https://doi.org/10.1177/0162243915608948

Zorina, A., Bélanger, F., Kumar, N., & Clegg, S. (2021). Watchers, aatched, and watching in the digital age: Reconceptualization of information technology monitoring as complex action nets. *Organization Science*. https://doi.org/10.1287/orsc.2021.1435

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75–89. https://doi.org/10.1057/jit.2015.5

Zuboff, S. (2021, January 29). Opinion | The coup we are not talking about. *The New York Times.* https://www.nytimes.com/2021/01/29/opinion/sunday/facebook-surveillance-society-technology.html

Humanities

Digital Whores Doing Pay-Per-View Chores

Deconstructing the Gentrification of Online Sexuality Through an Analysis of *OnlyFans*

Lisa Philippo

Supervisor
Dr. Erinc Salor (AUC)
Reader
Dr. Lara Mazurski (AUC)

Photographer: Daria Roman

Abstract

OnlyFans facilitates a digital marketplace to share and buy sexual content in a decreasing space for online sexuality after the enactment of the FOSTA/SESTA legislations. Over the course of the COVID-19 pandemic, the platform quickly rose in popularity and soon included profiles by various celebrities who are publicly praised for their presence on the sexual platform. This shows a stark contrast to the stigmatized place of sex workers in public discourse, which suggests a certain distinction between performers on *OnlyFans* and sex workers. To analyse this difference, this paper studies whether *OnlyFans* has changed the stigmatized perception of sex work, and how this relates to the identity and perceived agency of the performer. Through an analysis of sex work discourse, performers on *OnlyFans* are related to sex workers though discursive notions of agency, identity and stigma. Subsequently, a discourse analysis of digital sex work and NetPorn shows how *OnlyFans* compares with other digital platforms through interpenetration, mediated intimacy and aestheticism. The results of this analysis establish that although the popularity of *OnlyFans* may have broadened the public perception of sex work, the interpenetrative aspects of the platform that facilitated its popularity enlarge the disparity between influencers and sex workers - who are continuously excluded from participating in the (digital) marketplace. This demonstrates that the analysis of *OnlyFans* is illustrative of a larger digital gentrification, which perpetuates a glorified "influencer" aesthetic while slowly eliminating the digital presence of marginalized bodies and sexualities.

Keywords: *Sex work* – OnlyFans – *Digital Intimacy* – *Discourse Analysis* – *Online Platform*

Suggested citation:
Philippo, Lisa (2021). "*Digital Whores Doing Pay-Per-View Chores: Deconstructing the Gentrification of Online Sexuality Through an Analysis of* OnlyFans". AUC Undergraduate Journal of Liberal Arts and Sciences, vol. 16, no.1, pp. 96-116

Contents

1 Introduction

> *On that Demon Time, she might start a OnlyFans*
>
> *Big B and that B stand for bands*
>
> *If you wanna see some real ass, baby, here's your chance*
>
> *I say, left cheek, right cheek, drop it low, then swang*
>
> (Beyoncé on Meghan Thee Stallion's 'Savage' remix)

"*OnlyFans* Changed Sex Work Forever" stated a *New York Times* headline in 2019 (Bernstein). The subscription-based platform received growing media attention, quickly rising to pop-culture ubiquity when singer Beyoncé noted the platform in her remix of the most popular song of 2020 (Trust). While *OnlyFans* originally primarily facilitated sex workers, the popular platform now supports over one million content creators, including a variety of celebrities and influencers. *OnlyFans* is a digital platform "revolutionizing creator and fan connections", where performers create a profile behind a paywall, to which fans subscribe to see their content and pay for exclusive images and private messages, as explained on their website ("How it works"). While the *OnlyFans's* advertisement does not specify the type of performer, the minimized restrictions on content behind the paywall allow sex workers to post pornographic and explicit imagery that most digital platforms have banned after the FOSTA/SESTA[1] legislations were put into law (Pilipets and Paasonen 2; Pezutto 42). *OnlyFans* facilitates online sexuality in a diminishing space for sex workers both on- and offline, as anti-prostitution activists recently celebrated that the COVID-19 pandemic "prevents men from paying for sex" and renamed it the "abolitionist virus" (Döring 2773). After support for sex workers was largely "forgotten" in political responses to the pandemic, recent research reports that many 'in-person' sex workers have transferred their work to online platforms during the pandemic, *OnlyFans* being the predominant choice (Platt 9; Brouwers and Herrmann 13; Callander et al. 93). Sex workers shared their presence on *OnlyFans* through social media such as *Instagram* and *Twitter* which "are often the first point of contact through which aspiring performers learn about and get in touch with other performers" (De Jesus 81; Pezzutto 45). This public advertisement for *OnlyFans* coincided with growing numbers of global unemployment due to the COVID-19 pandemic, which encouraged recently unemployed people outside of the sex industry to join the platform. By the end of 2020, *The Guardian* reported that "everyone and their mom" was on *OnlyFans* (Lehmiller et al. 298; Boseley). A few months ago, *OnlyFans* announced that the platform now boasts over a hundred million users and paid over one billion US dollars to its performers that range from 'moms' and influencers to sex workers. Celebrated by popular media outlets such as *The New York Times* and *The Guardian*, *OnlyFans* appears to subdue much of the stigma "ubiquitous in sex work" which often discredits and excludes sex workers from public discourses (Weitzer 717). In order to investigate the distinctive attributes that allow *OnlyFans* to omit this discrediting stigma and 'change sex work forever', this research explores whether the platform has changed the stigmatized perception of sex work, and how this relates to the identity and perceived agency of the performer.

This study applies discourse analysis to investigate the divergent ways in which sex workers as well as their profession are discussed in the contrasting and overlapping discourses that collectively produce the perception of sex work. The analysis of sex work discourse poses an "open challenge" to the contemporary distinction between illicit and accepted sexual behaviour. Jane Scoular explains that sex worker narratives "offer important counter-hegemonic insights" that expose discursive "oppositions between erotic/affective activity and economic life" (347). This paper shows that *OnlyFans* is a relevant case study through which the contemporary perception of sex work can be analysed, as a critical analysis of the platform offers insight into the underlying tensions of agency, authenticity, intimacy, digitalization and intersectionality that are at play in sex work discourse.

Recently, researchers have made note of the "understudied intersection between relational work and technology" and urged future studies to connect "platform affordances to emotional and relational strategies" (Hair 209), while others mark the necessity of identifying "sex workers' dependence" on digital platforms (Brouwers and Herrmann 13). In the past five years, a growing body of research has recognized the increasing influence of digitally mediated platforms for intimate relations (Andreassen et al.; Hair; Paasonen; Swords), but none paid specific attention to sex work on these platforms. Similarly, digital sex work within discourses of Net-Porn (Jacobs et al.; Paasonen; Stardust; Miller-Young) and webcamming (Henry and Farvid; Jones) is widely researched but omits the interpenetrative aspects of modern digital platforms that rely on connections to hegemonic social media such as *Instagram* and *Twitter*. Although the research by Cunningham et al. on 'Commercial sex, digital spaces and working online' incorporates an analysis of digital platforms besides their quantitative research on sex work, the article was written in 2017 and describes adult platforms and forms of communication for sex workers that no longer exist after FOSTA/SESTA. This shows that most research on digital sex work is either focused on webcamming and porn, or is outdated. Although 2017 may not seem that long ago for other fields of research, the FOSTA/SESTA legislations heavily impacted the environment for digital sexuality, making research

[1] 'Allow States and Victims to Fight Online Sex Trafficking Act' (FOSTA) and 'Stop Enabling Sex Traffickers Act' (SESTA) is a legislative package of legislations enacted by the U.S. House of Representatives and the U.S Senate in 2018. A full explanation and discussion of their implications will be included further in the text.

on this topic before their implementation largely unapplicable to the current study. Recognizing the implementation of the FOSTA/SESTA legislations and the global spread of COVID-19, it cannot be ignored that recent developments have heavily impacted sex workers and influenced many aspects of the work on which previous research relied. Thus, the relevance of this research lies not only in its analysis of the relationality between a novel platform for digital intimacy and sex workers, but also in its unprecedented position of simultaneously reviewing digital sex work in light of FOSTA/SESTA and digital intimacy in light of the pandemic, while relating both to an intimate platform with unparalleled popularity.

In the establishment of a clear academic base to build research upon, the first chapter provides the necessary comprehensive framework by working through notions of stigma and illicit sexualities. Subsequently, a brief history of sex work discourse is included, as the aforementioned stigmatized discourse is exemplified and put into historical perspective. Likewise, more contemporary discourses that posit sex work as empowering and fuelled by agency are compared after which the chapter closes with a case study that examines whether stigma and agency in sex work are experienced equally. In the second chapter, this framework is applied to digital sex work, exploring aestheticism, authenticity, and intimacy in a developing online environment. The chapter goes on to consider the ways in which intersectionality defines a divergent experience in digital sex work, analysing the effects of FOSTA/SESTA on digital sexuality and its subsequent influence on *OnlyFans*. Finally, the discussion of the findings from this analysis shows that although the platform may have broadened the public perception of digital sex work, *OnlyFans* has not 'changed sex work forever' for everyone. The intermediate and interpenetrative qualities of the platform that initially enabled its growth directly restrict an inclusive and equitable representation of bodies and sexualities on *OnlyFans*, instead supporting a digital gentrification of online sexuality as the platform replaces sex workers for the more profitable influencers. The research concludes by observing the limitations of the present study and identifying future directions for further analysis.

2 Whores Versus Women: Sex Workers in Society

Sex work is often declared to be 'the oldest profession' (Ray 45; Henry and Farvid 113). However, the way in which it is acknowledged and discussed continuously changes. This chapter discusses the various aspects that are at play in the discourse of sex work and sex workers to establish a clear theoretical framework on which the research is built. After introducing the stigmatized discourse that works to exclude sex workers and demarcate them from a population, this chapter will analyse how this position has been affirmed through history. Following this, the contemporary discourse on sex work is explored, as this chapter moves towards a working definition of sex work as such. This definition then provides a base from which divergent aspects of sex work are recognized, such as notions of agency and identity. Having established a clear understanding of the most prominent discourses that produce our understanding of sex work, *OnlyFans* is situated in this shifting landscape of definitions. The chapter ends with a discussion of the present tensions and complexities through a case study of Bella Thorne.

2.1 Whore Theory

"The prostitute is the prototype of the stigmatized woman. She is both named and dishonoured by the word 'whore'", explains Gail Pheterson (65). This is echoed by Ronald Weitzer, who notes that "stigma is ubiquitous in sex work" (717). As visible in Pheterson's explanation, the stigma related to the profession of prostitution instantly reproduces the identity of the worker through the frame of that stigma. "Stigma, then", notes Erving Goffman, refers to "an attribute that is deeply discrediting" (3). The attribute of a stigmatized profession thus discredits the worker and dishonours their identity, which is visible in discourses of sex work that perpetuate stigmatized narratives of either exploitation or promiscuity. This positions the worker as a helpless slave on the one hand, and an overtly sexual agent on the other (Smith 347). Both discourses discredit the identity of the sex worker, the first by denying their agency and the latter by condemning their morality. However, Goffman continues to explain that although stigma functions as a discrediting attribute, it should be seen as "a language of relationships, not attributes" (3). The discrediting identity of stigma only works within a certain relationality; "a special kind of relationship between attribute and stereotype" that makes certain stigmatized attributes only discrediting to a specific group of people, while there remain "important attributes that almost everywhere in our society are discrediting" (Goffman 4). Although Pheterson and Weitzer's earlier statements imply that prostitution is one of these universally discrediting attributes, Pheterson does note that the 'whore stigma' related to prostitution is also used to discredit women outside the profession (65). Jill Nagle confirms this, explaining that whores are "something that women are not only supposed to not be, but also, not be *mistaken* for" (5 emphases in the original). Women are expected to ensure that they are never perceived as whores, a responsibility that weighs in with every minor decision, as being perceived as a whore implies an instant degradation to "every dimension of dishonour" (Pheterson 89). This responsibility works as a restrictive burden to those who carry it, and is not only placed on those who sell sex (Pheterson 66). In fact, it can be quite easy for women to be perceived as a whore. As understood from Pheterson's text, one only has to wear too few clothes, wear the wrong clothes, have too much sex, have sex with the wrong person, have sex at the wrong time, speak at the

wrong time, be enthusiastic, easy, different, or show any other slight exposure of illicit behaviour that contradicts the chaste expectations of women. "Signs of indecency or unrestraint in women are thought to reflect inherent qualities which distinguish the character of whores from non-whores. Whores are expected to be '*different kinds of women*'" explains Pheterson (86 emphasis in original). In this way, illicit connotations attached to the job of 'selling sex' are transformed into an infinite set of characteristics that transfer these same connotations to any woman labelled as a whore. Thus, 'whore' suffices as a synonym for any woman that is unchaste or exposing illicit (sexual) behaviour. This effectively produces a policing of sex, explained by Michel Foucault in *The History of Sexuality* (vol. 1) as "the necessity of regulating sex through useful and public discourses" (25). Any sexuality that does not fit a discourse of utility is disregarded from the public discourse and alienated as illegitimate. In this way, a clear distinction between illicit and legitimate sexualities is enforced through a constant redefinition and reproduction of what sexuality is and should be. All sexualities that do not fit this produced norm are instantly reduced to illicit sexualities.

Similarly, women who do not uphold expected standards of decency are reduced to whores. Foucault clarifies the historical connection between these two groups:

> If it was truly necessary to make room for illegitimate sexualities, it was reasoned, let them take their infernal mischief elsewhere: to a place where they could be reintegrated, if not in the circuits of production, at least in those of profit. The brothel would be a place of tolerance [...] Only in those places would untrammelled sex have a right to (safely insularized) forms of reality. (4)

In this quote, Foucault traces the pairing between 'abnormal' sexualities and prostitution back to the 18th century, when a necessity to organize the population instilled a focus on the regulation of sexuality. It is important to acknowledge that the primary idea of Foucault's argument here entails that –contrary to popular belief– sex was never repressed (6). Instead, it was regulated and redefined through discourses of utility, aiming for "an ordered maximization of collective and individual forces" (25). Thus, instead of sex being dismissed and ignored, the 18th and 19th century showed a pattern of institutionalization of sex, following an eruption of interest in theory on sexuality (48). It is impossible to think of sex work outside these institutional discourses, as they have produced and shaped sex work in (Western) society[2].

[2]It is important to acknowledge that this study primarily focuses on a 'Western' cultural frame. Future studies are recommended to review this research from a globally different point of view, as the historical narrative provided here is not universal.

2.2 A Brief History of Sex Work Discourse

> The history of modern prostitution control offers a dynamic perspective on the private lives of women as well as the public functioning of medicine, patriarchy and the nation state and emphasizes the need to understand how gender and sexuality are interrelated inextricably to race, cultural diversity and economic circumstances. (Guy qtd. in Scoular 343)

The institutional production of a differentiation between whores and 'normal women' can be traced back through several faculties. Perhaps most notorious is Cesare Lombroso's 1893 work *La Donna Delinquente*, where he identified significant physical differences between whores and women according to his theory of anthropological criminology, defining prostitution and criminality as genetically determined. Although Lambroso's theory was largely refuted by the beginning of the 20th century, *La Donna Delinquente* remained the primary text for the field of female criminology until the 1970s (Rafter and Gibson 4). This demonstrates how - in line with Foucault's institutional production of discourses - the necessity to regulate and categorize sex work produces a new academic discourse of utility: criminal anthropology. Subsequently, the impact of this discourse, however refuted, clearly informs the discursive understanding of sex work in the late 20th century and onwards as both inherently related to crime and unmistakably visible in the female body. This provides a base for the inherent difference between whores and women, and 'educates' women to restrict their physicality to prevent being perceived as a whore (Pheterson 86).

A similar genealogy is noticeable in the reproduction of sex work through medical discourses. "Historically, even before sexual transmission of disease was understood, women prostitutes were blamed for epidemics such as The Plague", explains Pheterson (68). Prostitutes were so deeply connected to dirtiness and disease that this provided a reason to medicalize them through discourses of physical and mental defects (Petro 165; Beloso, "Queer" 148; Pheterson 70; Fofana 5). With regard to physical defects, Brooke Beloso explains how the decrease of the leprosy epidemic by the end of the Middle Ages opened the stage for a fear of venereal diseases, which were, just as The Plague, directly linked to prostitutes (146). However, in contrast to the distanced but inclusive approach to individuals with the plague in earlier decades, leprosy was approached through systems of expulsion and confinement, which served as a precedent for the response to venereal diseases (Beloso, "Queer" 145). This resulted in international orders to expel and transport prostitutes from cities to places of confinement (Beloso, "Queer" 146). *In History of Madness*, Foucault explains that "it was only under the influence of the world of confinement in the seventeenth century that venereal disease became detached to some extent from this medical context, and like madness entered a space of social and moral exclusion" (8). While in

comparison to leprosy, a variety of medical 'cures' for venereal diseases evolved relatively quickly and led to its sufferers being removed from places of confinement to medical institutions. However, prostitutes remained excluded (Beloso, "Queer" 147). Following this, prostitutes inhabited places of confinement or 'working houses' alongside "the 'debauched', 'imbecile', 'prodigal', 'infirm', 'of unsound mind', 'libertine' 'insane" (...) "and all were cast into the same abstract dishonour" (Foucault 81). Here, the shift from medical confinement to 'moral exclusion' is visible, as prostitution is produced through discourses of insanity and immorality. As a result of this reproduction of prostitution through the image of The Madman, prostitutes who were hospitalized due to venereal diseases up until the late 19th century were 'cured' through whippings and moral purging as it was argued that the disease's "directional mode of transmission reflected prevailing *attitudes* rather than any bacteriologic reality" (Foucault 85; Pheterson 68, emphasis added). It is not surprising that if the contemporary understanding of sex work is established through these discourses of dishonour and immorality, the stigmatized perception of sex workers that lies at their foundation remains ubiquitous. Accordingly, sex workers are still among the first to recieve blame in the case of a pandemic, as in recent years. This was unmistakably noticeable during the AIDS and COVID-19 pandemics, which both showed a strong moral accusation of sex workers (Pheterson 68; Döring 2773; Fofana 5). It is important to recognize that each of these early understandings and discursive reproductions of sex work directly informs and creates the contemporary discussion of sex work. By tracing the historical discourse of sex work, early connections between sex workers and criminality, physical appearance, dirtiness, disease and morality are visible. Accordingly, Foucault explains that the essential part is:

> Not to determine whether one says yes or no to sex (. . .) but to account for the fact that it is spoken about, to discover who does the speaking, the positions and viewpoints from which they speak, the institutions which prompt people to speak about it and which store and distribute the things that are said. (*History of Sexuality* 11)

With this in mind, more contemporary discourses of sex work can be considered in order to relate the productive and archival institutions currently at work to their histories. The contemporary debate of sex work is notably characterized by the feminist 'sex wars' of the late twentieth century which divided the feminist movement over debates regarding pornography, sadomasochism and female sexuality in general (Scoular 345). Beginning around the 1970s with 'radical' feminists such as Barry, MacKinnon, and Dworkin, sex work was defined in a discourse of slavery and exploitation which posed sex workers as 'helpless victims' in an inherently subjugating and oppressive patriarchal system (Scoular 345; Petro 155; Kempadoo 43). Radical feminists argue that sex should only take place in the private sphere, under ideal and equal circumstances, for "when sex is not explicitly treated as genuine human interaction, it dehumanizes the experience and thereby dominates women" (Barry 28). Where radical feminists see all sex work as the 'commodification of women', 'sex positive' or 'liberal' feminists believe in a differentiation between sex out of love and sexuality in other roles and purposes. Liberal feminist academics such as Rubin, Willis, and Doezema, argue that the potential harm experienced by sex workers is not inherent to the work, but rather a result of "specific regimes of criminalization and denigration that serve to marginalize and oppress sex workers while constraining and distorting sex work's radical potential" (Zatz 298). This argument aligns with the outcomes of a variety of qualitative interviews with sex workers from different continents who state social denigration and police violence as their primary problems (Buijs; Sanders; DMSC; Gülcür and Ilkkaracan). Henry and Farvid explain that the radical feminist's understanding of sex work is grounded in their academic approach that focusses "less on the individual and seek[s] to focus on dismantling oppressive social structures, arguing that individual 'choices' are inherently constrained rather than 'free' within a patriarchal context" (114). However, this argument effectively removes all agency from sex workers and instead defines them as powerless subjects to an inescapable system. In their resistance to the dominant patriarchal system, radical feminists only see resolution in a radical dismantling of that system, which effectively posits women's choices within this patriarchal system as irrelevant. Remembering Lambroso's theory of 'born prostitutes', the anti-patriarchal narrative of radical feminists appears to have internalized a similar deterministic pessimism. Where Lambroso argues that prostitutes lack a free choice for the profession as they are simply subjected to their genetic heritage, sex radicals discredit sex worker's autonomy in a similar way, stating that they are always already 'constrained within a patriarchal context'. Additionally, it is interesting to note that both Lombroso and the abolitionist movement that evolved from radical feminists' theory urge for the criminalization of sex work.

Judith Butler warns radical feminists for a nostalgic imaginary of a utopian "before" patriarchy and female subordination, which can be "problematic when it constrains the future to materialize an idealized notion of the past or when it supports, even inadvertently, the reification of a precultural sphere of the authentic feminine" ("Gender" 47). Chasing this idealized imaginary excluded and alienated sex workers from the feminist movement, which fragmented over the unsolvable disputes (Sheiner 39; Queen 128; Butler, "Gender" 48). While radical feminists retained their idealistic aim to at once dismantle the patriarchal system, liberal feminists sought inclusionary strategies based on a "recognition that sex work is an integral part of the global economy and is deeply embedded in, and cannot easily be disassembled from, many women's everyday lives, strategies, and identities" (Kempadoo 44). Although sex-positive sex workers are

in turn often criticized for their idealized picture of sex work as empowering, it cannot be overlooked that sex workers' agency is shaped by such gravely different histories and experiences based on their divergent genders, cultures and ethnicities, which "cannot be simply reduced to one monolithic explanation of violence" (Kempadoo 28). I will return to the culturally divergent experience of agency later in this chapter, however, I would first like to highlight another of Kempadoo's arguments against the abolitionist 'modern slavery' narrative. Sex work theory that "start[s] from a 'victim approach'", explains Kempadoo, is thoroughly based on "the belief that they can rehabilitate men who use prostitutes and can "rescue" or "save" women in a missionary fashion" (43). This provides an interesting insight into the paradox between the 'modern slavery' narrative of sex work abolitionists and the colonial 'white saviourism' in which it is rooted (Beloso, "Sex" 65).

2.3 Sex Work Identities

Sex worker, author, and activist Carol Leigh – or Scarlot the Harlot – brought the term 'sex work' into use in the 1970s (225). She initially created the term to remove the focus of the conversation regarding prostitutes from their identity as such, which was often a problem with terms containing negative connotations such as 'prostitute' or 'whore' (Leigh 226). Instead, the term 'sex work' solely emphasizes the occupation of the person. Now, more than thirty years after Leigh's rearticulation of the conversation, sex work has become the dominant term in academic discourse and has provided a strong tool for sex worker's self-identification (Berg, "Working" 693). Viewing sex work as work allows its entry into a discourse of labour, moving away from a focus on crime, disease, danger and even sexuality. Likewise, Leigh's term paved the way for the primary mantra of the decriminalization movement: 'Sex Work Is Work' (*Red Umbrella Fund*).

Another advantage of the term 'sex work' is its lack of specificity. It can include any sexual act that in some way may be seen as work, such as stripping, making porn or being a phone sex operator. While this inclusivity enlarges the spectrum of the sex work community, it also encourages people to reconsider the imagined boundaries of sex work. Sex work is often defined as 'the exchange of sexual services for resources, goods or money' (Henry and Farvid 113; Petro 156; Scoular 345). This broad definition includes many types of sex work and pays attention to the different things it can be exchanged for. However, what counts as a sexual service remains unclear. Does this definition focus on the sexual nature of the action that the worker carries out, or is it up to the client, viewer or receiver to decide what a sexual action is, and what is worth paying for (Henry and Farvid 118)? Sex-work activist Heather Berg proposes that instead of this focus on either sexual services or economic gain, sex work is far more defined by the sociocultural understanding it provides to the workers ("Working" 692). Although the 'sex work as work' narrative reinstalls the focus of prostitution on its values as labour and this, in turn, diverges this attention away from the former hegemonic discourse of stigma, Berg argues that it remains important to retain these specific elements that distinguish sex work from other work:

> It is the work of free exchange between equals, the dignity of a living earned and a heady blend of both self-sacrifice and fulfilling escape from the drudgery of a nine-to-five-job. It is sometimes work that is barely work at all, but instead a performance for the innate self for which the lucky just happen to be paid. ("Working" 693)

A valuable aspect of this definition is the sole interest in the experience of sex work for the worker. This enables an interesting conversation concerning whether we can, or should, regard sexual exploitation as part of sex work. With regard to Berg's definition, this is not the case. The position taken in this definition thus also responds to abolitionist arguments that define all sex work as inherently exploitive (Leigh 230). Following Berg, sex work cannot be equated to exploitation, since the foundation of its definition is based on a free choice.

"Freedom consists in the ability to autonomously 'choose' one's desires no matter how illiberal they may be", states Saba Mahmood (12). Accordingly, any action is not inherently oppressive if the subject has the free autonomy to choose it. Although sex work may seem illiberal or harmful to some, this sense of autonomy and agency is essential for sex workers (Barry 33; Jones, "Motivations" 103). Besides the necessary agency in the 'free exchange' of the work, many sex workers cite obtaining agency as a primary motivation behind their specific choice for the work (Jones, "Motivations" 103; Miller-Young, "Hypersexuality" 221; Petro 161). Next to the freedom of 'nine-to-five' hours, which I will analyse later, this experience of agency is largely based on the performative qualities that are inherent to sex work. This part in Berg's definition is unmistakably built off Butler's work on performativity which Mahmood also connotes to, as she elaborates on how Butler imagined the "power of performatives" to oppose normative structures as a direct act of agency (21). Sex work is a continuous performance of sexuality, where intimacy is 'acted out' as a commodity in direct opposition to normative structures which confine intimacy and sexuality in (heteronormative) discourses of utility and decency such as the nuclear family.

Where Berg identifies sex work as a 'performance for the innate self' she relates to the personal agency that can be gained in the free decision for a counter-hegemonic action. A similar experience is found in interviews by Angela Jones who describes that sex workers are "reconciling pleasure with work" by playing with normative notions of both work and sexuality as a sex worker states "I get to have fun and play with myself" ("Motivations" 102, 96). This "play labour" does not entirely overturn or redefine any norms, but pre-

vents alienation and helps the worker to regain agency in both their work and sexuality (Miller-Young 225). Jones uses John Holloway's theory of Crack Capitalism to describe the de-alienating effects of rediscovering pleasure in work ("Motivations" 90). By challenging capitalism in individual and quotidian ways, for example by finding labour in leisure activities, cracks may appear in the system. Comparing this system to a frozen lake, the cracks may never break through all of the ice, as they would not dismantle the entire system, "but they do help the worker to feel dignity again" (91).

While all workers can experience the alienating effects of labour in a capitalist system, finding quotidian ways to reconcile pleasure with work can be especially valuable to groups that are historically marginalized by capitalism. Recognizing the ever-present influence of intersectionality, race plays a crucial role in sex work theory (Kempadoo 28). Miller-Young confirms this, emphasizing that "illicit eroticism offers black women who are marginalized in the formal labor market ways to survive and prosper" (224). Black women are inherently connected to sex work through racist binaries that perceive both sex and race as "dark, mysterious, and dirty. Both are judged unchaste and thereby unfit for public life" (Pheterson 71). Where white women need to restrict their sexuality in an attempt not to be perceived as a whore, black women are declared unchaste by nature and instantly equated to whores (Kempadoo 40; Fofana 5). As a result of this, black female bodies are both hypersexualized through their perceived inherent relationality to 'dirty sexuality' and deemed disposable by their inability to ever become a good, chaste woman (Kempadoo 40; Nagle 5; Miller-Young, "Hypersexuality" 226; Pheterson 76). In line with Foucault, these illicit sexualities do not find a place in discourses of utility or the public sphere (25). Thus, the marginalized position of black women often forces them to find or create labour outside the public and inside the private, intimate sphere. Here, sex work may provide a valuable opportunity as "black sex workers, while facing multiple axes of discrimination and harm, also employ hypersexuality and illicit eroticism to achieve mobility, erotic autonomy, and self-care" (Miller-Young, "Hypersexuality" 224).

Performing sex work as a method to reclaim erotic autonomy is described by various groups that are excluded as illicit sexualities, such as trans people. (Royalle and Sundahl 160). Trans, Latina sex worker Cecilia Gentili explains that "for trans people, who are constantly told that they are not valid, that they should not exist and are not attractive, it can be very empowering to feel wanted and receive this validation on your own terms" (Farvid et al.). Here, the goal is not to overthrow the system of heteronormativity [3], it is about finding agency by reclaiming a marginalized identity and turning it into value: "you can fetishize the shit out of me, as long as you pay me" (Gentili in Farvid et al.).

[3]The performativity of sex work plays a crucial role in LGBTQIA+ communities, in both its stigmatizing and constricting aspects, as in its liberating potentialities. There is much more to explore here, but this is outside of the scope of the present work. For more information on trans and non-binary experiences in digital sex work see Jones "Cumming" and Pezzutto. For representation of sex work in queer discourse see Beloso "Queer", and Pendleton, among others.

2.4 *OnlyFans* in Sex Work Discourse

The rise of *OnlyFans* in early 2020 and its subsequent ubiquity in popular media is stated to be largely due to its concurrence with the beginning of the COVID-19 pandemic (Brouwers and Hermann 13). In response to the pandemic, multiple national governments enacted a nationwide 'lockdown', limiting social contact and urging people to stay at home and transfer their day-to-day lives to an online environment (Grasso et al. 9). Following a significant decrease in in-person labor, many people lost their jobs or acquired an entirely different working routine. This abrupt shift incited a re-imagination of labor, as people were forced to search for new and different ways of gaining income in the private, intimate sphere. This search was not only incited by economic necessity, but also to maintain social contact and a daily structure to prevent alienation, avoiding what Foucault would characterize as "confinement (. . .) which created alienation" ("History of Madness" 80).

Part-time, self-employed work from home has long been "consigned to the working classes and/or women", who were historically forced to find labor outside of the "standard employment model" (Rand 42). A primary example of this self-employed 'gig'-work is sex work, which, as earlier discussed, has long provided agency and income to groups that were excluded from the public economy. Recently, the lower-class image of the gig-economy has shifted towards the profile of a "creative entrepreneur", who uses digital platforms to maintain relationships, advertise their profile and individually create an income (Hair 197; Rand 42). Despite increased recognition of gig-work, Rand describes that digital, platform-based sex work was still ignored from the discourse of gig-work, as digital sex work was "devalued as non-work, 'bad' work or illegitimate work due to the social stigma experienced in varying degrees by sex workers" (Rand 44). However, as working from home became the standard, and it became necessary to find new ways of gaining income, the national lockdowns were found to further "improve the value and prestige" of platform-organized gig-work.

It appears that this shift has also inspired an improved view of the possibilities of digital sex work (Grasso et al. 9; Drolet). In response to restrictive regulations of physical sex work in the lockdown, many sex workers were quick to make the shift from physical sex work to online services such as *OnlyFans*, promoting their presence on the platform through *Twitter* and *Instagram* and thereby helping the platform gain popularity (Pezzutto 45; Döring 2773; Callander et al. 93). The consistent mention of *OnlyFans* on social media inspired those looking for new sources of income to "take the plunge", resulting in a seventy-five percent increase of subscriptions to

the platform within six months of the first COVID-19 peak in January 2019 (Drolet; *OnlyFans*). However, the enormous increase of new members to *OnlyFans* may present a risk to the sex workers that initially incited its growth. The result of this is painstakingly visible in the case of Bella Thorne, a former Disney actress who joined the platform to sell her fans explicit, nude images for two hundred US dollars. Despite this promise, she was not nude in the photos and thousands of subscribers demanded a refund of their tips and subscriptions from *OnlyFans* (Russon). In response to this, Russon reports that *OnlyFans* placed a limit on the amount that performers can ask for a photo or video, and limited the amount a fan can tip a performer.

Additionally, *OnlyFans* changed their payment system from weekly pay-outs to a single payment per month. Thorne claimed she earned two million dollars in two days, while many performers with a smaller following are heavily impacted by *OnlyFans'* policy change (Russon). The limit to tips and material damages performers who created their business model around high-rate customized content, while the new structure of monthly payments has a large influence on performers whose livelihood's depend on their incomes from *OnlyFans*. In response to the criticism, Thorne explained on *Twitter* that she joined the platform to "remove the stigma on sex work" by "bringing a mainstream face to it" (@bellathorne). Through this comment, Thorne aligns herself with a long history of academics, celebrities and journalists that argue to support sex workers by speaking about them, instead of amplifying their voices (Nagle 3). Gentili emphasises the problems of this narrative, explaining that besides the damage that celebrities do to sex work by joining a conversation they are not part of "they have also done extreme damage by using the damage that they did in their favour" (Farvid et al.). Although Thorne stated that she was in conversation with representatives of *OnlyFans*, the new structure remained and Thorne retained her money and her *OnlyFans* account.

Regardless of whether it was meant to remove a stigma or solely for personal gain, Thorne's position on sex work illuminates the divergent ways in which sex work is performed and experienced based on intersectionality. Thorne never self-identified as a sex worker, despite joining and still earning from an account on *OnlyFans*. It is interesting to note that through her supposed aim to support sex workers, she distances herself from the other sex workers on the platform, emphasising the sex worker as the 'other'. Subsequently, this position establishes the sex worker as the victim, who 'needs' Thorne's help to remove the stigma. Thorne's distinction between her 'mainstream face' and sex workers also visualizes the grey area of sex work that *OnlyFans* represents (@bellathorne). Although the platform gained its popularity and initial financial growth through sex workers, the increasingly overwhelming presence of celebrities on the platform might redefine who the platform represents. While Leigh's definition of 'sex work' includes any type of sexual labour and allows every individual to choose whether they prefer to use this term, the breadth of the term also leaves space for polarizing differentiations within the profession. This clearly distinguishes those who momentarily appropriate the profession for their personal gain without experiencing its legacy of stigma, and the sex workers who do not have that choice.

3 Power to the Platform: The Precarity of Online Sexuality

Why do people pay for a subscription-based platform of sexual media when there is an abundance of free pornography all over the internet? To answer that question, this chapter explores digital sexuality in the contemporary interconnectivity of the Internet. Where the first chapter created a discursive and theoretical framework of sex work in which *OnlyFans* can be located and compared, this chapter relates *OnlyFans* to contemporary theory on digital platforms and online sexuality. Through discourses of authenticity, intimacy and labour politics, this chapter aims to outline the potentiality of digital sex work, while remaining critical of its inclusive realities.

3.1 NetPorn Aesthetics

Susanna Paasonen describes a turn from online pornography to NetPorn halfway through the first decade of the 21st century, the same moment Andreassen et al. note a general shift of living "*in* media, rather than *with* media" (*Introduction*, emphases in the original; 1299). Specific to online platforms and digital communication networks, NetPorn entails

> The blurred boundaries of porn producers and consumers, the proliferation of independent and alternative pornographies, as well as the expansion of technological possibilities brought forth by digital tools, platforms and networked communications. (Paasonen 1298)

The intertwined networks of communication in the contemporary Internet landscape enable a diminished experience of the boundaries established by the screen in between viewer and performer, allowing everyone to be both consumer and producer, entertainer and enjoyer, through what Paasonen describes as "peer-to-peer exchanges" (1302). This creates a collaborative experience where sexual material can directly be requested online and catered to by performers, or shared with a larger audience according to their choice. This direct link between performer and consumer, either through personal collaboration or self-uploaded material, omits the necessity for a production company as a third party in the creation of pornography (1299). Platforms that allow pornographic content without the restrictions and expectations of

a production company facilitate a much lower threshold for 'deviant' bodies, sexualities and identities to get into pornography, allowing them infinitely more creative freedom (Jacobs et al. 2).

Self-uploaded, 'amateur' material outside of large porn production companies also caters to a demand for authenticity (Jones, "Scroll" 787). Paasonen observes that where 'professional' porn is signified by its high-quality cameras and lighting, glossy, 'perfect' bodies, "amateurism implies the opposite, [as] it is also coded in terms of spontaneous, 'more truthful, and less-manufactured representations'" (1303). Recognizing a paradox in the apparently inherent connection between amateurism and authenticity, Paasonen explains that "home media seem to involve modes of double-thinking: the images are selective and knowingly produced, yet function as evidence (. . .) to verify sexual arousal and climax" (1303). Amateur pornography thus provides a sense of visual realism, which answers a need for the 'verification' of arousal. Although the viewer is aware that the imaginary is manufactured and that the sexual situations are deliberately created, the lower-quality aesthetics still connote a verification of sexual "truth" (1303).

A result of this notion of the 'truthful' appearance of sexual pleasure is the association between amateur porn and 'ethical porn', where "a system of aesthetics is evoked as a form of ethics" (Attwood qtd. in Paasonen 1300). While the vocabulary of ethical porn as such enforces an array of discursive problems which is further exposed by Mondin, the entire implication of 'truth' in pornographic production creates a problematic binary that poses 'good porn', which is less produced and more ethical, alternative, collaborative, realistic and artistic as a direct opposite to 'bad porn' which, signified by its high-quality productions implicates non-realistic or untruthful sexuality, and would immediately lack all the earlier mentioned aspects (Stardust 15). Interestingly, this shows how the discourse of the seemingly positive turn to NetPorn has exposed a new example of the returning differentiation between good and bad or illicit sexuality, where new discourses of utility continuously are defined through a strong denouncement of its supposed opposite (Beloso, "Queer" 154). Yet again, this emphasizes how "all acts of subversion are a product of the terms of violence that they seek to oppose" (Mahmood 21).

The idealization of authentic and 'true' sexuality in NetPorn discourse can be explained by the "desire to see someone doing something because they like to be seen" (Paasonen 1303). Similarly, Jones noted that the intercommunicative tendencies of NetPorn create a space where it feels like the performers are "real people who have turned on their webcams" where they "invite clients into their real bedrooms, and are genuinely enjoying themselves" ("Cumming" 2). Crucial to this experience are the truthful *intentions* of the performer, enforcing "the expectation that performers act out of love and not money" (Henry and Farvid 118; Stardust 19). It seems absurd to discern that after decades of discourse victimizing sex workers for being exploited because they earned money from having sex, discourses of authenticity now require the workers to do their job for free to prove they are not being exploited.

3.2 Digital Intimacy

Zahra Stardust describes that the "fantasy of genuine sexuality" is driven by "an entitled consumer demand for personalised intimacy – for free access to not only performer's bodies but their intimate and authentic pleasures" (19). This demand appears to move away from the voyeuristic tendencies of (amateur) porn and move towards an entirely digital intimacy (Jacobs 2). Characterized by the possibilities and intercommunicative imperatives of the contemporary Internet, Andreassen et al. define digital intimacy alongside social media, as it can "allow people to express and share what matters to them, and both encourage personalized connection and interactivity". The authors explain that the contemporary digital landscape facilitates a shift for both interpersonal connections and intimacy, from the private sphere to the (digital) public sphere (*Introduction*). While NetPorn generally encompasses collaborative and communicative sexual production, digital intimacy requires much more personalized labour from workers, who are expected to engage in intimate, personal relationships with their clients. Therefore, Henry and Farvid identify digital sex work as highly emotional labour. This allows sex work to be analyzed through the framework of emotional labour, which is usually set apart for other (female-dominated) professions such as nursing and hospitality (117). This comparison can be productive, on the one hand, to recognize the similarities in the serving and nurturing nature of both sex work and hospitality, and on the other, to pay more attention to the emotionally exhaustive qualities of both professions (122).

Digital labor blurs the boundaries between private and public, between work and home, making it much harder to distinguish work and leisure. While physical sex workers may rent a space to invite their client over, digital sex work is usually performed from home. Similarly, physical sex work often ends when either the client or the worker leaves, while digital sex work can continue all day, as a digital sex worker Kristy reports: "For me I don't feel I get a break. My phone does not stop. I get customers texting me all the time" (qtd. in Rand 50). In addition to using a continuous online presence to respond to clients' messages, digital sex workers are expected to keep their followers updated on their personal and professional lives, create and upkeep genuine online identities, while still performing "the emotional labour of convincing consumers that one enjoys the work but also a process of managing consumer access to it" (Stardust 20). In this way, digital sex workers make a living from the paradoxical performance of a 'manufactured genuine identity', which is central to the emotional labour they are paid for, while simul-

taneously "monetizing new forms of identity" and managing this income to generate more clients (Stardust 24).

The performance of sexual identity in sex work shows an exaggeration of everyday gender performances as described by Butler in *Bodies that Matter*, where they emphasize that "what determines the effect of realness is the ability to compel belief, to produce naturalized effect" (5). Although a manufactured performance of identity and sexuality can be experienced as "a resistance strategy" that provides sex workers with agency and control in the workplace, sex workers often invest part of "their personality into their working identity, rather than constructing an entirely fake working identity" to create a realistic, natural performance (Sanders 337, 334). The relationality between intimacy and identity is clearly explained by Andreassen et al., who note that "contemporary intimacy is often analysed as an illustration of how an individual (. . .) investigates their self-reflection in relation to close relationships and narratives of identity" (*Introduction*). Likewise, as intimacy has become a marker for self-reflection and identity, the manufactured performance of intimacy for digital sex workers also influences their self-perception of identity. As a result of the obscured distinctions between digital sex workers' professional and personal identities, Rand describes the experience of a 'presence bleed', where digital sex workers risk losing touch with their identity and face burn-outs as work and non-work become inseparable (49; Stardust 24).

3.3 Class in the Digital Marketplace

The emotionally heavy aspects of sex work can differ per type of sexual labour, argues Bernstein, as she states that the sexual labour of middle-class sex workers is more likely to include "emotionally engaged conversation as well as a diversity of sexual activities" (482). Although the diversity of sex workers between "high school graduates and multiple degree holders" had been repeatedly documented among sex workers, Ray states that "before the advent of the Internet", the media still portrayed sex workers as "either high-class escorts (. . .) or downtrodden, drug-addicted streetwalkers who were usually women of color" (Dudash 89; 49). Bernstein proposes that the growing presence of the middle-class sex worker is strongly facilitated by "new technologies of sexual exchange", which enable "the curious woman to explore her working options" without any commitments (474; Ray 49). Over 15 years after Ray and Bernstein's prediction, Scoular et al. confirm that 'internet-enabled sex workers' represent a different profile than other sex workers, "in terms of generally higher average age, gender balance, and level of qualifications" (215). This can be partially explained by the fact that 'service economies' such as digital intimacy valorize "rigid ideas about the employees' aesthetic presentation of self" (Jones, "Scroll" 790). Therefore, the digital intimacy market glorifies the ideal white, thin, educated, able-bodied, cisgender performance, while "workers who fall outside that ideal overwhelmingly make less money, have fewer work opportunities, and encounter poorer working conditions" (Berg, "Scene" 170). Although NetPorn was hailed for its promotion of "gender fluid entrepreneurism" or "voluptuous pleasure", the digital intimacy market continues to avoid nonconforming bodies (Jacobs et al. 2). Simultaneous to Bernstein and Ray in the early twenty-first century, Miller-Young proposed that the representation of different bodies and sexualities in NetPorn may force a change in the larger industry, as queer and trans directors and production companies work to "subvert or queer the heterosexist, racist capitalism of cybercultural space" (Miller-Young, "Sexy" 208). However, Jones later shows that the 'cis male gaze' is so deeply embedded in mainstream porn, that contemporary inclusion of marginalized communities often still perpetuates a narrative of "white cis male privilege and dominance" ("Cumming" 12). While growing initiatives provide new opportunities in porn that aim to subvert colonial narratives and provide a fair pay to performers, these productions are still in the minority and the assumption that all workers have the liberty to choose in which productions they participate is simply another performance of white privilege: "If they could afford to turn it down, you wouldn't see as many companies offering them work for less money" (Berg "Scene" 170).

3.4 FOSTA/SESTA: The 'Pornpurge' of the Internet

Where minorities are structurally last to benefit from new advancements, they are also first to encounter the danger of decreasing social spaces. Recently, this has become painstakingly clear with the new adoption of the FOSTA/SESTA laws. In response to the growing market for digital sex work, the abolitionist movement followed with a "call for greater regulation of information technology which is seen as facilitating the 'industrialized commodification of women'" (Scoular et al. 217). While police forces initially lagged behind on these changes and retained their focus for trafficking regulation on street prostitution, this heavily changed in 2018 with the adoption of the 'Fight Online Sex Trafficking Act' (FOSTA) and the 'Stop Enabling Sex Traffickers Act' (SESTA) (Scoular et al. 221). These laws hold online platforms accountable for all content that is shared and generated by users of the platform, making the companies behind the platforms directly responsible for any on or offline 'unlawful actions' related to content posted on the platform (Tripp 219). Accordingly, the laws criminalize any "'interactive computer service' with 'the intent to promote or facilitate the prostitution of another person'" (Campbell 1557). The package containing the FOSTA/SESTA laws was quickly brought into action by U.S. president Trump after a vigorous 'stop modern slavery' media campaign by the abolitionist movement, which re-installed sex trafficking on the top of the political agenda (Campbell et al. 1556; Scoular et al. 238). As a result of the social pressure behind the fast production of the law, the restrictions are not specifically articulated,

leaving 'prostitution' undefined (Tripp 236). Consequentially, there is no distinction made between consensual sex workers and victims of trafficking, as they are all merged under prostitution and reproduced through a discourse of slavery (Mia 239). Additionally, FOSTA/SESTA also impacted sex educators, academics and artists among others from all over the world, as platforms such as *Instagram*, *Tumblr*, *Facebook*, and the search engine Google, chose to completely censor all 'sexual imagery' and 'sexual speech' through 'automated filters' (Pilipets and Paasonen 7). Of course, this decision was not made because these profiles were genuinely advertising human trafficking, but "because policing them against the outside possibility that they might was just too hard" (Tripp 237). In this way, the effects of FOSTA/SESTA are visible in every part of the digital sphere as academic sexology groups are deleted, female nipples are restricted, educational images are censored and sex workers' PayPal accounts are frozen (Pilipets and Paasonen 13; Mia 239; Tripp 238).

Although FOSTA/SESTA was initially installed to safeguard sex workers and prevent sexual exploitation, various researchers emphasize the explicit danger that sex workers face as a result of the laws. Tripp describes how digital advertisement enabled sex workers to find clients through the internet, instead of self-advertising on the streets. Additionally, this function allowed sex workers to screen potential clients more extensively before meeting them (229). Campbell et al. warned of the implications of FOSTA/SESTA on "networking, peer support and information sharing" among sex workers, "which are vital for screening and wider safety" (1557). Ray exemplified this through their emphasis of the importance of shared 'blacklists' among sex workers which catalogue potentially dangerous clients. Likewise, Tichenor argues that their removal "created an environment where people are having to crawl back to brothels", to ensure their safety (58; 107).

3.5 *OnlyFans* in the Digital Space

NetPorn heavily expanded the possibilities of digital intimacy and inspired a new, more individual type of online sex work outside of high-produced pornography. However, these possibilities are not equally attainable to all sex workers, and for those for whom the work is their sole source of income, maintaining a digital sex work profile can be emotionally draining and exhausting. "Through the lens of class analysis one begins to see the mutually imbricated privilege and oppression", explains Beloso ("Sex" 64). Systems of white privilege and white saviorism are deeply intertwined and extremely visible in both the sex-positive and abolitionist sides of sex work discourse. Where the constant connections to slavery and a presumed need to save women from traffickers in the FOSTA/SESTA legislations are clear examples of white saviorism, the early NetPorn discourse structurally assumes that the technological advancements of the internet will benefit all sex workers equally. For example, while the (white) editors of the academic NetPorn reader praised its potentiality in the feminist pornographic movement, they overlooked the disparate ways in which many of these possibilities are only viable options for white, middle-class sex workers. In response to the broader trend of unacknowledged white privilege in feminist pornographies, Jones urges:

> Before we applaud new digital technologies for the ostensibly agentic and lucrative potential spaces they create for women, let us realize that these benefits are clearly contingent on the race and nationality of the model. ("Scroll" 795)

With regard to *OnlyFans*, a similar pattern is visible. Mickey Mod, the creative director of 'ethical' porn company Kink.com stated that the increasing popularity of *OnlyFans* among (former) porn performers has enlightened many of the possibilities of NetPorn as it allowed them "to choose how they worked, and directly find out from their fans what they wanted and provide to that" (Farvid et al.). Mod applauded that *OnlyFans* enables performers to "create their own terms and boundaries and start from there", which would "force porn producers to take these changes seriously" (Farvid et al.). However, Jones notes that "given the value placed on gender normativity, whiteness, and able-bodiedness in the camming market, people whose bodies do not conform to these standards will not earn decent wages" ("Cumming" 9). In order to 'start from your own boundaries' or 'force change', performers first need to find themselves in a position of power, which is structurally not the case for non-conforming bodies who have to fight up to a "market with both less work to go around and more people willing to do it" (Berg, "Scene" 164). This is especially difficult in light of the extreme popularity of *OnlyFans* during the lockdown, which inspired so many people to join that this flooded the market and minimized the chances to successfully make a living through the platform (Boseley; Brouwers and Herrmann 13).

As of this writing, the thirty highest-earning profiles are all owned by celebrities who, just like Bella Thorne, moved to the platform with the support of their followers on other social media. Most of these celebrities do not identify themselves as sex workers, or do not post 'sexual' content at all. Twenty-three of these best-earning *OnlyFans* performers are cisgender white women ("Blog"). While the popularity of the platform inspired many to reconsider how they imagine labor and find agency in a performance of the self, most of the incomes from *OnlyFans* accounts do not go to either sex workers or workers affected by the COVID-19 pandemic, but are instead received by celebrities who use the platform as an additional source of income. This can be traced back to the platform affordances of *OnlyFans*, which avoids 'scrolling through pages' of performers, as profiles can only be found through a direct link instead of through a search bar. As online sex work is per definition aesthetic labor, bodies that do not conform to the normalized expectations of perceived

attraction are usually pushed away into the corners of online platforms, because "racialized bodies end up with lower bodily capital in the sexual marketplace", explains Jones in the aptly-named article *For Black Models Scroll Down* (791). The digital structure of *OnlyFans* appears to combat this, allowing each performer the same chance of building their profile and attracting followers as there is no list of profiles to 'scroll' through. However, as the only way to share the direct link is through other social media, performers' dependence on the interpenetration of other platforms such as *Twitter* and *Instagram* instantly removes this apparent equality (Pezzuto 45; Swords 530). *Instagram* is centered around an aestheticism that idolizes the thin, white, female body, the result of which is that it is much easier for women who fit these standards to grow a large following, which can then be directed to their *OnlyFans* profile (Jones, "Cumming" 9; Paasonen, "Infrastructures"). This is directly visible in the list of the highest earners on *OnlyFans*, who all were able to transfer their large amount of pre-existing followers on other social media to *OnlyFans*.

The interpenetration of *Instagram* and *Twitter* within *OnlyFans* becomes an even larger problem in light of the FOSTA/SESTA legislations, after which any mention of sex work or related words within digital platforms is heavily restricted and often leads to a directly deleted account without a warning. Sex workers have repeatedly stated that they are being unfairly locked out of their accounts without posting any 'forbidden' material, and argue that their profiles are deleted based on the presumption of their profession (Musto et al. 11; Pezzutto 44). Another large concern for sex workers on digital platforms is being 'shadow-banned', this occurs when their account is not removed, but simply made invisible for potential new followers. This blocks the sex worker from "an important source of marketing, income, and socialising", furthermore, as it is difficult to verify whether the account is actually shadow-banned, sex workers experience "serious anger and anxiety" (Pezzutto 43). Although these restrictions are entirely automatically operated, Pilipets and Paasonen explain that "algorithmic decisions about what should be removed and what should stay on social media platforms reenact a biased social system connected to sexuality, surveillance, and computation", as the differentiation between 'acceptable' and 'unacceptable' bodies and sexualities "is never neutral nor uncontested" (11). In this way, thin, white and generally conforming bodies have a much lower chance of being banned through algorithmic decisions for posting similar content compared to bodies that do not conform and thus stand out to the algorithm.

There are more aspects where *OnlyFans* aligns its digital structure with other social media, such as in the discursive notion of authenticity for its outstanding performers. In contrast to authenticity in NetPorn discourse that is related to amateurism and lower quality, homemade imagery, performers on *OnlyFans* can earn a 'badge of verification' if they have "at least 10,000 followers on [their] Twitter or Instagram Accounts" ("Help"). Only verified performers have the chance to be 'featured' by *OnlyFans*, where their profile is shared and promoted throughout the social accounts of *OnlyFans* and in their blog. Accordingly, only the accounts that follow the guidelines of Instagram can receive this mark of authenticity and be supported by *OnlyFans* and thus, gain more fans. This further enlarges the gap between influencers and celebrities that already have a big following on social media, and sex workers who are repeatedly banned from the same platforms – often for posting similar content – as the former are supported and promoted by *OnlyFans*, and the latter have to work much harder to achieve the same level of support. This discrepancy is unmistakably clear in the 'How it works' page on *OnlyFans*, which provides a tool to calculate how much one can earn on the site, depending on the amount of Instagram and Twitter followers the performer already has ("How it works"). This exemplifies how the 'algorithmic structures of acceptability' that resulted from the FOSTA/SESTA legislations reshape online sexuality. Sex workers are not allowed to state their profession on social media. Furthermore, as semi-sexual content is heavily restricted as well as dependent on the body by which it is performed, it is incredibly difficult to navigate the abundance of platform guidelines while also attracting followers to direct to their *OnlyFans*. Subsequently, the strict policing of sexuality on Instagram also influences performers' 'authenticity' on *OnlyFans*, as they can only receive the verification badge by connecting with a high-profile Instagram account.

Similarly, the interpenetration of *Twitter* and *Instagram* on *OnlyFans* and its strong distinction from other Netporn platforms is also visible in the digital aesthetics of *OnlyFans*. As the social possibilities of *OnlyFans* coincided so perfectly with an increased demand for sexual intimacy in the pandemic, the platform quickly became representative of the social and technological possibilities of NetPorn to sex workers and porn performers who found a source of autonomy in using the platform (Jesus 81; Farvid et al.). However, this autonomy is only notable outside of the NetPorn discourse, which had long recognized the autonomy of individual production. Interestingly, while *OnlyFans* does provide intimate contact with fans, collaborative performances, real-time uploads, performer independence and many more aspects that are characteristic of NetPorn, the aesthetic tendencies of *OnlyFans* are completely different. As earlier explained, NetPorn was hailed for its 'do-it-yourself' amateur aesthetic, which was defined by its contrast to the high-quality production of 'mainstream' pornography. The interface of *OnlyFans* in turn, also shows a highly stylized aesthetic (see fig. 1 in appendix), which is most comparable to the interface of Instagram (see fig. 2 in appendix) and shows a stark contrast to other NetPorn platforms such as Chaturbate (see fig. 3 in appendix)[4]. In light of the aforementioned problematic binary distinction between NetPorn and 'mainstream' pornography, where the first is seen as amateur, alternative, independent and ethical, and

[4]A visual analysis of the stylistic similarities between *OnlyFans* and *Instagram* can be insightful but is not within the scope of this work.

the latter as professional, commercial and exploitive, *OnlyFans* presents an interesting combination. The platform facilitates a variety of self-uploaded material, which can be uploaded by amateurs or professionals, and can contain anything from filmed orgies and domination to bikini photos and baking tutorials. The performer – or 'creator', as *OnlyFans* calls its members – can style their profile and manufacture their digital image as realistically or imaginary as they prefer, while fulfilling the demand for personalized digital intimacy through private messages and requests. With regard to the expectation that amateur sex workers should not expect a monetary exchange if they perform their sexuality for their own pleasure, *OnlyFans* leaves it to its creators to decide whether they want to charge their fans.

3.6 Profiling the Platform: Digital Gentrification and the Influencer Class

Although the popularity of *OnlyFans* is often related to the lockdowns during the COVID-19 pandemic, the platform appears to fulfil a range of other demands. Produced by and simultaneously taking part in the production of a changing digital landscape, *OnlyFans* presents itself as an answer to the intersection of an ever-growing influencer-culture where fans demand more from their idols, a restriction of online sexuality during a rise in demand for online intimacy, and a mass loss of jobs coinciding with a growing gig-economy. Despite the very ambitious and promising image painted by the platform, the analysis in this chapter proposes that some of these divisions may remain mutually exclusive. The interpenetration of *Instagram* and *Twitter* in *OnlyFans* helped promote the platform to a larger audience, while simultaneously encumbering the admission and advertisement of sex workers and people with 'lower body capital' in Instagram's aestheticism. By only allowing individual profiles and content behind a paywall, *OnlyFans* avoids the restrictions on sexual imagery as imposed through the FOSTA/SESTA legislations and provides a certain 'safe space' for online sexuality. However, the suggestion of 'physical' sex work is heavily policed through equally generalized algorithms, as an increasing amount of sex workers report that their *OnlyFans* accounts were blocked and their incomes were frozen without breaking the guidelines (Dickinson). In contrast to Bella Thorne, who was immediately reported to join conversations with the management of *OnlyFans* following the repercussions of her content, smaller profiles are ignored by the 'customer service' and have no way to regain access to their account (Russon; Dickinson). For a considerable portion of the sex workers who joined *OnlyFans* after restrictions on their physical jobs due to lockdown measures, the revenue generated through the platform is their primary source of income and the unexpected suspension heavily impacts their lives (Callander et al 96; Mia 239; Jones, "Cumming" 9; Dickinson). In this way, the effects of *OnlyFans*' automated policing further separates sex workers from influencers on the platform. Although representatives of the platform stated that "without question, *OnlyFans* is one of the most inclusive social platforms, and [their] progressive policies towards content creation enables the success of adult content creators without discrimination", this celebration of adult content is not visible in the public profile, advertisements or social media presence of the platform, which all focus on celebrities and influencers (Dickinson; "Blog"). The disparate experiences of sex workers and celebrities on the platform emphasize that "not everybody can sell their sex equally", as privilege and oppression are two sides of the same coin, with the result that "privilege along one axis can mitigate oppression along another, even within the sex industry itself" (Beloso, "Sex" 64).

Celebrities joining *OnlyFans* appear to largely avoid the social expulsion still experienced by many sex workers as, instead of being discredited, their voices are amplified while their presence on the platform is celebrated and shared through social media, a practice usually embedded in risk for non-celebrity performers on *OnlyFans*. Pezzutto describes the higher social class of celebrities and sex workers with a high number of followers on *Instagram*and *Twitter* as "a hierarchy within the industry", which produces and reinforces a growing class disparity (47). A high number of followers on social media means a higher chance to get verified, which helps to get featured on other channels and is often followed by lucrative sponsored partnerships that lead to more fans on *OnlyFans*(Berg, "Scene" 168). Moreover, these additions often provide a better and more stable income than the platform itself does (Pezzutto 46).

> If it is true that sexuality is the set of effects produced in bodies, behaviors, and social relations by a certain deployment deriving from a complex political technology, one has to admit that this deployment does not operate in symmetrical fashion with respect to the social classes, and consequently, that it does not produce the same effects in them. (Foucault, *History of Sexuality* 127)

Bodies that are more sexualized, such as black and trans bodies, as well as illicit sexual behaviors and non-heterosexual social relations, perform deviant attributes of sexuality which, as earlier explained, would be relocated to the brothel where, in circuits of profit "untrammelled sex [would] have a right to (safely insularized) forms of reality" (*History of Sexuality* 4). Deviant sexualities, whether physical, behavioral or relational, are thus discursively produced through the stigmatized image of the prostitute which inherently relates those sexualities to sex work. In this way, the social perception of an *OnlyFans* profile is dependent on the hegemonic attributes of the performer. Accordingly, this further clarifies why Bella Thorne cannot "remove the stigma on sex work" by "bringing a mainstream face to it", as the stigma "does not produce the same effects" on her, through her inherent disassociation with the stereotype (@bellathorne; Foucault

127). Where Foucault notes an asymmetrical effect of the political production of sexualities concerning social classes, the interpenetrative qualities of digital platforms reproduce this hierarchical distinction through amounts of followers, establishing not the bourgeoisie, but influencers and celebrities as the digital upper-class. The composition of social classes may have changed, but the distribution remains the same, as one percent of *OnlyFans* earns thirty-three percent of all income on the platform, closely resembling how the top one percent of households in the world own a near forty percent of global wealth (Zucman 111).

Through the increasing interdependence of platforms on the interpenetrative hegemony of *Instagram*, *Facebook*, and *Twitter* in the contemporary Internet landscape, the normative aestheticism and purified sexualities that these enforce, regulate the presence and subsequent perception of bodies and sexualities throughout the Internet. As the effects of FOSTA/SESTA continue to unfold, influencers prove a more secure and rewarding group of performers to *OnlyFans* as a company than sex workers, as visible in the featured accounts and influencer-directed advertising of the platform. While an increasing amount of sex workers are banned from *OnlyFans* without explanation, a digital 'gentrification' of the platform is deployed, replacing the sex worker presence on the platform with the more profitable and upper-class influencers.

4 Conclusion

This paper analyzed whether *OnlyFans* changed the stigmatized perception of sex work, and measured how this relates to the identity of the performer through an extensive discourse analysis. Throughout the analysis, a variety of sex worker narratives were celebrated for their counter-hegemonic possibilities, yet after a critical analysis, the case study of *OnlyFans* shows that both sex positive and abolitionist narratives imagine digital sex work from a position of privilege as centuries-old perceptions of sex work still define its discourse. While the performance of sex work can playfully act out heteronormative perceptions of sexuality, this is not as much a form of resistance as it is often a requirement for deviant sexualities to successfully join the digital marketplace. This is exemplified in Gentili's disregard of the constant fetishization of her body because she is paid: although she allows this on her own terms, a profitable performance is still expected to follow an oppressive patriarchal narrative, as fetishization is deployed from a position of power. A similar tension is visible between the initial celebration of NetPorn in the early twenty-first century and the digital representation of bodies and sexualities fifteen years later. NetPorn was anticipated to queer the cybercultural space by relocating the directive power from big production companies to individual performers which would promote non-conforming bodies and sexualities and provide performers with a worthy pay. However, marginalized performers often lack the liberty to maintain their boundaries and choose a queer production company or an appropriate salary, as these discussions can only be held from a position of power outside of a discourse where their bodies are inherently regarded as disposable. Likewise, self-directing digital content allows sex workers to further perfect their manufactured digital identity as a performative resistance strategy, but also results in unattainable pressure and a presence bleed.

These findings do not imply that sex work cannot be experienced as agentic and liberating, but instead propose a more nuanced approach to its counter-hegemonic potentiality. Developments in digital sex work such as *OnlyFans* can certainly work towards a less stigmatized perception of sex work. However, this thesis demonstrates that stigma is relative to the individual it is related to, instead of a set of universally discrediting attributes. While the stigma of sex work can disgrace women who do not participate in the work based on their connection to a stereotypically related attribute such as promiscuity, the stigma does not have the same discrediting effects on those who do not match the stereotype. *OnlyFans* proudly advertises that 'anybody' can join the platform, which may abate the stigmatized idea that sex workers are different kinds of women, yet the distinction between 'whores and women' remains persistent within *OnlyFans*, as visible in Bella Thorne's tweets and the disparate treatment of *OnlyFans* to sex workers and influencers.

Thus, the results of this analysis establish that although *OnlyFans* has broadened the perception of sex work, introducing more people to the possibilities of digital sex work as it lowered the threshold to join an adult platform, '*OnlyFans* has not changed sex work forever', as it does not influence the stigmatized perception of women who were already discredited as 'whores' in any way, and instead continues to perpetuate this exclusion. While it is no longer necessary to 'scroll down' for black performers on *OnlyFans*, this does not change the perception of black and trans bodies as disposable and of lower bodily capital. This is especially visible in the response of social platforms to the FOSTA/SESTA legislations, where this discursive depreciation of non-conforming sexualities was – albeit unknowingly – directly deployed in the regulatory algorithms that valorize thin and white bodies and digitally ostracized any slight suggestion of deviant sexualities.

This analysis of *OnlyFans* serves as an example for a larger pattern of digital gentrification in relation to the interpenetration of hegemonic social media that enforce increasingly strict regulations on sexuality and glorify a very specific and exclusionary aesthetic. This paradigm is a strong starting point for future research, both on digital sexuality and platform-organized sex work. While the present research analyzed a multiplicity of overlapping and interconnected discourses to situate *OnlyFans* in its academic environment and illuminate its distinctive features, this also serves as a limitation, as this method obstructed an exhaustive investigation of a single fea-

ture of the platform. Additionally, sex work discourse is heavily influenced by the stigmatized assumption that sex work is female labor, which is why this research primarily focused on female sex work performances. Similarly, the present study restricted the analysis to Western discourses of sex work. Future studies are therefore recommended to further include male and non-binary sex worker narratives, and approach the subject from a divergent cultural framework. Nevertheless, this thesis provides an illuminating insight into the interplay between the relationality of stigma in perspectives on sex work, and digital affordances such as interpenetrative platforms and digital gentrification that established both the rise of *OnlyFans* into popular culture and the elimination of illicit sexualities off the platform.

Works Cited

"About". *OnlyFans*. `OnlyFans.com/about`, Accessed on 8 May 2021.

Andreassen, Rikke, et al.. Introduction. *Mediated Intimacies: Connectivities, Relationalities and Proximities*. Routledge, 2017, doi:10.4324/9781315208589.

Barry, Kathleen. *The Prostitution of Sexuality*. NYU Press, 1995. JSTOR, `www.jstor.org/stable/j.ctt9qg779`. Accessed on 25 May 2021.

Beloso, Brooke Meredith. "Sex, Work, and the Feminist Erasure of Class." *Signs*, vol. 38, no. 1, 2012, pp. 47–70, doi:10.1086/665808.

—. "Queer Theory, Sex Work, and Foucault's Unreason." *Foucault Studies*, no. 23, 2017, pp. 141–66, doi:10.22439/fs.v0i0.5345.

Berg, Heather. "'A Scene Is Just a Marketing Tool': Alternative Income Streams in Porn's Gig Economy." *Porn Studies*, vol. 3, no. 2, 2016, pp. 160–74, doi:10.1080/23268743.2016.1184478.

—. "Working for Love Loving for Work." *Feminist Studies*, vol. 40, no. 3, 2014, pp. 693–721, `link.gale.com/apps/doc/A390997199/AONE?u=amst&sid=AONE&xid=2df6c3c0`.

Bernstein, Elizabeth. "Sex Work for the Middle Classes." *Sexualities*, vol. 10, no. 4, 2007, pp. 473–88, doi:10.1177/1363460707080984.

"Blog." *OnlyFans*. `https://blog.OnlyFans.com`, Accessed on May 11, 2021.

Boseley, Matilda. "'Everyone and Their Mum Is On It': *OnlyFans* Booms in Popularity During the Pandemic." *The Guardian*, 22 December 2020, www.theguardian, Accessed on 24 April 2021.

Brouwers, Lilith, and Tess Herrmann. "'We Have Advised Sex Workers to Simply Choose Other Options'—The Response of Adult Service Websites to COVID-19." *Social Sciences*, vol. 9, no. 10, 2020, pp. 1-15, doi:10.3390/socsci9100181.

Buijs, Laurens. *De Amsterdamse Prostitutiemonitor*. Amsterdam Sociaal, 2010.

Butler, Judith. *Gender Trouble*. Routledge, 1999, doi:10.4324/9780203824979.

—. *Bodies that Matter: on the Discursive Limits of Sex*. Routledge, 1993, doi:10.4324/9780203828274.

Callander, Denton, et al. "Investigating the Effects of COVID-19 on Global Male Sex Work Populations: A Longitudinal Study of Digital Data." *Sexually Transmitted Infections*, vol. 97, no. 2, 2021, pp. 93–98, doi:10.1136/sextrans-2020-054550.

Crenshaw, Kimberlé. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum*, 1998, pp. 139–67, `chicagounbound.uchicago.edu/uclf/vol1989/iss1/8`.

Dickinson, EJ. "Sex Workers Built *OnlyFans*. Now They Say They're Getting Kicked Off." *Rolling Stone*, 18 May 2020, `https://www.rollingstone.com/culture/culture-features/OnlyFans-sex-workers-porn-creators-999881/` Accessed on 17 May 2021.

Döring, Nicola. "How Is the COVID-19 Pandemic Affecting Our Sexualities? An Overview of the Current Media Narratives and Research Hypotheses." *Archives of Sexual Behavior*, vol. 49, no. 8, Springer, 1 Nov. 2020, pp. 2765–78, doi:10.1007/s10508-020-01790-z.

Drolet, Gabriella. "Sex Work Comes Home: More of Us Are Making and Watching Sexual Performances Online Now. Fewer of Us are Paying." *The New York Times*, 10 April, 2020, `https://www.nytimes.com/2020/04/10/style/camsoda-OnlyFans-streaming-sex-coronavirus.html`, Accessed on 6 May 2021.

Dudash, Tawnya. "Peepshow Feminism". *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 89–118, doi:10.4324/9780203700655.

Farvid, Pani, et al. Sex Work in a Pandemic: Legislation and Technology. *The New School*, 30 March 2021, `https://livestream.com/thenewschool/sexworkinpandemic`. webinar.

Fofana, Mariam. "Decolonising Global Health in the Time of COVID-19." *Global Public Health*, 2020, pp. 1-12, doi:10.1080/17441692.2020.1864754.

Foucault, Michel. *History of Madness*. Translated by Jonathan Murphy and Jean Khalfa, edited by Jean Khalfa, Routledge, 2006, doi:10.4324/9780203642603.

—. *The History of Sexuality vol. 1:The Will to Knowledge*. Translated by Robert Hurley, Penguin Books, 1998, doi:10.1086/ahr/84.4.1020.

Goffman, Erving. "Stigma and Social Identity." *Stigma: Notes on the Management of Spoiled Identity*, 1974, pp. 1–19, doi:10.4324/9780203793343-4.

Grasso, Maria, et al. "The Impact of the Coronavirus Crisis on European Societies. What Have we Learnt and Where Do We Go From Here? –Introduction to the COVID Volume." *European Societies*, 2021, pp. 1-31, doi:10.1080/14616696.2020.1869283.

Gülçür, Leyla, and Pınar İlkkaracan. "The "Natasha" experience: Migrant Sex Workers From the Former Soviet Union and Eastern Europe in Turkey." *Women's Studies International Forum*, vol. 25, no. 4, 2002, pp. 411-421, doi:10.1016/s0277-5395(02)00278-9.

"Help." *OnlyFans*. `https://OnlyFans.com/help`, Accessed on May 11, 2021.

Henry, Madeline, and Panteá Farvid. "'Always Hot, Always Live': Computer-Mediated Sex Work in the Era of 'Camming.'" *Women's Studies Journal*, vol. 31, no. 2, 2017, pp. 113-128, doi:10.18574/nyu/9781479842964.001.0001.

"How it works." *OnlyFans*. `https://OnlyFans.com/how`, Accessed on May 11, 2021.

Jacobs, Katrien, et al. Introduction. *C'Lick Me: A NetPorn Studies Reader*, edited by Katrien Jacobs, Marije Janssen and Matteo Pasquinelli, Institute of Network Cultures, 2007, pp. 45-68, `https://www.networkcultures.org/_uploads/24.pdf`.

Jones, Angela. "Cumming to a Screen near You: Transmasculine and Non-Binary People in the Camming Industry." *Porn Studies*, 2020, pp. 1–16, doi:10.1080/23268743.2020.1757498.

—. "For Black Models Scroll Down: Webcam Modeling and the Racialization of Erotic Labor." *Sexuality and Culture*, vol. 19, no. 4, 2015, pp. 776–99. doi:10.1007/s12119-015-9291-4.

—. "Global Motivations to Cam." *Camming: Money, Power, and Pleasure in the Sex Work Industry*, New York University Press, 2020, pp. 87–108, doi:10.18574/nyu/9781479842964.001.0001.

Kempadoo, Kamala. "Women of Color and the Global Sex Trade." *Meridians*, vol. 1, no. 2, 2001, pp. 28–51, doi:10..1215/15366936-1.2.28.

Lehmiller, Justin, et al. "Less Sex, but More Sexual Diversity: Changes in Sexual Behavior during the COVID-19 Coronavirus Pandemic." *Leisure Sciences*, vol. 43, no. 1–2, Routledge, 2021, pp. 295–304, doi:10.1080/01490400.2020.1774016.

Leigh, Carol. "Inventing Sex Work." *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 223-231, doi:10.4324/9780203700655.

Mahmood, Saba. "The Subject of Freedom." *Politics of Piety: The Islamic Revival and the Feminist Subject*. Revised edition, Princeton University Press, 2005, pp. 1-39, doi:10.7208/chicago/9780226248646.003.0026.

Mia, Valentina. "The Failures of SSESTA/FOSTA: a Sex Worker Manifesto." *Transgender Studies Quarterly*, vol. 7, no. 2, 2020, pp. 237–39, doi:10.1215/23289252-8143393.

Miller-Young, Mireille. "Putting Hypersexuality to Work: Black Women and Illicit Eroticism in Pornography." *Sexualities*, vol. 13, no. 2, 2010, pp. 219–235, doi:10.1177/1363460709359229.

—. "Sexy and Smart: Black Women and the Politics of Self-Authorship in NetPorn." *C'Lick Me: A NetPorn Studies Reader*, edited by Katrien Jacobs, Marije Janssen and Matteo Pasquinelli, Institute of Network Cultures, 2007, pp. 205–14, `https://www.networkcultures.org/_uploads/24.pdf`.

Mission & Principles. Red Umbrella Fund, `https://www.redumbrellafund.org/about-us/mission-principles`, Accessed on 6 May 2021.

Mondin, Alessandra. "Fair-trade porn+ niche markets+ feminist audience." *Porn Studies* vol. 1, 2014, pp. 189-192, doi:10.1080/23268743.2014.888251.

Nagle, Jill. Introduction. *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 1–15, doi:10.4324/9780203700655.

Paasonen, Susanna. "Infrastructures of Intimacy and the Deplatforming of Sex." *Berlin Institute for Cultural Inquiry*, organized by Peter Rehberg, Hanna Hamel, and Apostolos Lampropoulos, 11 February 2021, `https://www.ici-berlin.org/events/susanna-paasonen`, Online lecture.

—. "Labors of Love: Netporn, Web 2.0 and the Meanings of Amateurism." *New Media and Society*, vol. 12, no. 8, 2010, pp. 1297–312, doi:10.1177/1461444810362853.

Pendleton, Eva. "Love for Sale: Queering Heterosexuality." *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 73–82, doi:10.4324/9780203700655.

Petro, Melissa. "Selling Sex: Women's Participation in the Sex Industry." *Sex Work Matters*, edited by Melissa Hope Ditmore, Antonia Levy, and Alys Willman, Zed Books, 2010,

pp. 155–69, doi:10.5040/9781350222526.ch-009.

Pezzutto, Sophie. "From Porn Performer to Porntropreneur: Online Entrepreneurship, Social Media Branding, and Selfhood in Contemporary Trans Pornography." *About Gender*, 2019, pp. 30–60, doi:10.15167/2279-5057/AG2019.8.16.1106.

Pheterson, Gail. "The Whore Stigma: Crimes of Unchastity." *The Prostitution Prism*, Amsterdam University Press, 1996, pp. 65–89, doi:10.5117/9789053561850.

Pilipets, Elena, and Susanna Paasonen. "Nipples, Memes, and Algorithmic Failure: NSFW Critique of Tumblr Censorship." *New Media and Society*, 2020, pp. 1-22, doi:10.1177/1461444820979280.

Rafter, Nicole Hahn; Gibson, Mary. Editor's Introduction. *Criminal Woman, The Prostitute, and the Normal Woman* by Cesare Lombroso, 1893, Duke University Press, 2004, pp. 3-34, doi:10.1215/9780822385592-028.

Rand, Helen. "Challenging the Invisibility of Sex Work in Digital Labour Politics." *Feminist Review*, vol. 123, no. 1, Nov. 2019, pp. 40–55, doi:10.1177/0141778919879749.

Ray, Audacia. "Sex on the Open Market: Sex Workers Harness the Power of the Internet." *C'Lick Me: A NetPorn Studies Reader*, edited by Katrien Jacobs, Marije Janssen and Matteo Pasquinelli, Institute of Network Cultures, 2007, pp. 45-68, `https://www.networkcultures.org/_uploads/24.pdf`.

Royalle, Candida, and Sundahl, Debi. "First Ladies of Feminist Porn." *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 156–166, doi:10.4324/9780203700655.

Russon, Mary-Ann. "Bella Thorne, OnlyFans and the Battle over Monetising Content." *BBC*, 1 September 2020, `https://www.bbc.com/news/business-53979625`, Accessed on 6 May 2021.

Sanders, Teela. "'It's Just Acting': Sex Workers' Strategies for Capitalizing on Sexuality." *Gender, Work and Organization*, vol. 12, no. 4, 2005, pp. 319–42, doi:10.1111/j.1468-0432.2005.00276.x.

Scoular, Jane. "The 'Subject' of Prostitution: Interpreting the Discursive, Symbolic and Material Position of Sex/Work in Feminist Theory." *Feminist Theory*, vol. 5, no. 3, 2004, pp. 343–55, doi:10.1177/1464700104046983.

Scoular, Jane, et al. "Beyond the Gaze and Well beyond Wolfenden: The Practices and Rationalities of Regulating and Policing Sex Work in the Digital Age." *Journal of Law and Society*, vol. 46, no. 2, June 2019, pp. 211–39, doi:10.1111/jols.12155.

Sex Workers Manifesto. Durbar Mahila Samanwaya Committee (DMSC), India, 1997.

Sheiner, Marcy. "Odyssey of a Feminist Pornographer." *Whores and Other Feminists*, edited by Jill Nagle, Routledge, 1997, pp. 37–43, doi:10.4324/9780203700655.

Stardust, Zahra Zsuzsanna. "From Amateur Aesthetics to Intelligible Orgasms: Pornographic Authenticity and Precarious Labour in the Gig Economy." *About Gender*, vol. 8, no. 16, 2019, pp. 1–29, doi:10.15167/2279-5057/AG2019.8.16.1115.

Swords, Jon. "Interpenetration and Intermediation of Crowd-Patronage Platforms." *Information, Communication & Society*, vol. 23, no. 4, 2020, pp. 523-538, doi:10.1080/1369118x.2018.1521455.

Tichenor, Erin. "'I've Never Been So Exploited': The Consequences of FOSTA-SESTA in Aotearoa New Zealand." *Anti-Trafficking Review*, vol. 14, 2020, pp. 99–115, doi:10.14197/atr.201220147.

Trust, Gary. "Megan Thee Stallion & Beyonce's 'Savage' Surges to No. 1 on Billboard Hot 100." *Billboard*, 5 May 2020, www.billboard.

Zatz, Noah D. "Sex Work/Sex Act: Law, Labor, and Desire in Constructions of Prostitution." *Signs*: Journal of Women in Culture and Society, vol. 22, no. 2, 1997, pp. 277–308, doi:10.1086/495157.

Zucman, Gabriel. "Global Wealth Inequality." *Annual Review of Economics*, vol. 11, 2019, pp. 109–138, doi:10.1146/annurev-economics-080218-025852.

@BellaThorne. "PT1 Remove the stigma behind sex, sex work, and the negativity that surrounds the word SEX itself by bringing a mainstream face to it that's what I was trying to do, to help bring more faces to the site to create more revenue for content creators on the site." *Twitter*, 30 Aug. 2020, 12:24 a.m. `https://twitter.com/bellathorne/status/1299835299132891136`.

Appendix

Note: I chose to compare *OnlyFans* to *Chaturbate* because the platform is not only the most popular, but also mentioned in *Global Motivations to Cam* (Jones), *Cumming to a Screen Near You* (Jones), *The Failures of SESTA/FOSTA* (Mia) and *Always Hot, Always Live* (Henry and Farvid)

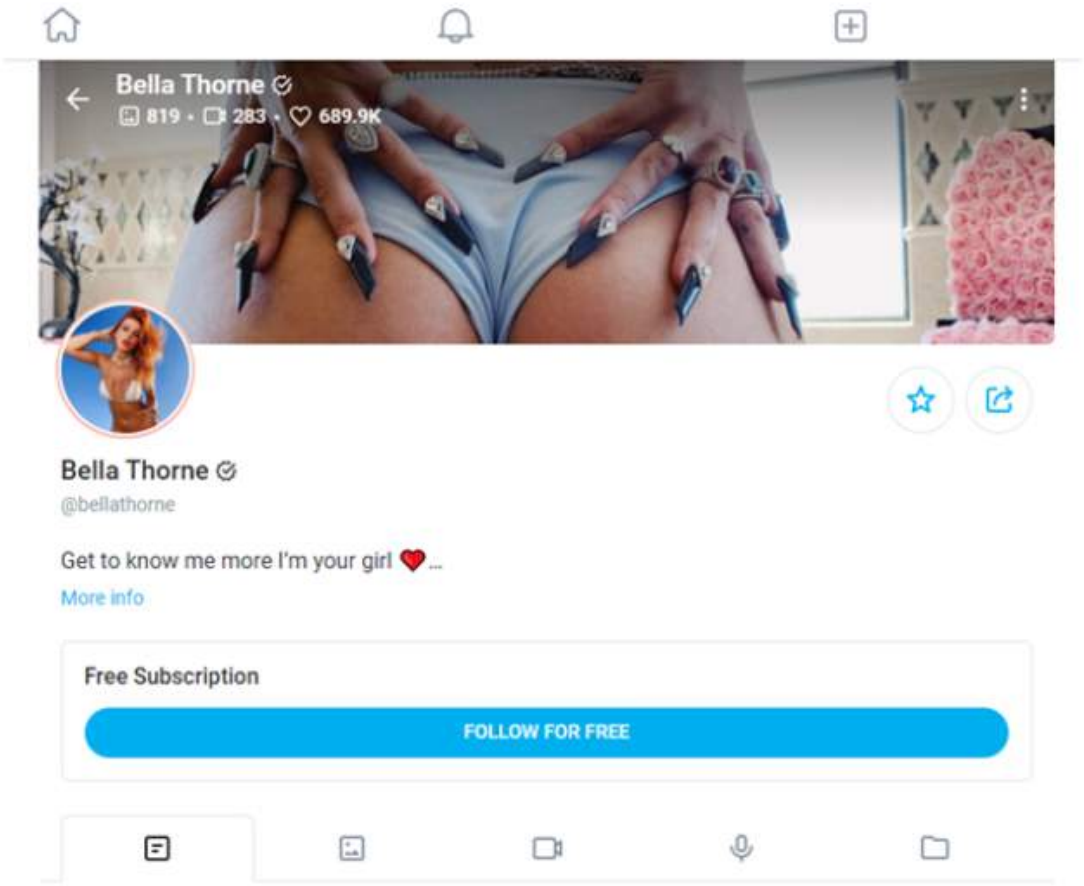

Fig 1. Philippo, Lisa. "Non-subscriber Interface of the *OnlyFans* Webpage of Bella Thorne", *OnlyFans*, May 11 2021, `www.OnlyFans.com/bellathorne`, screenshot.

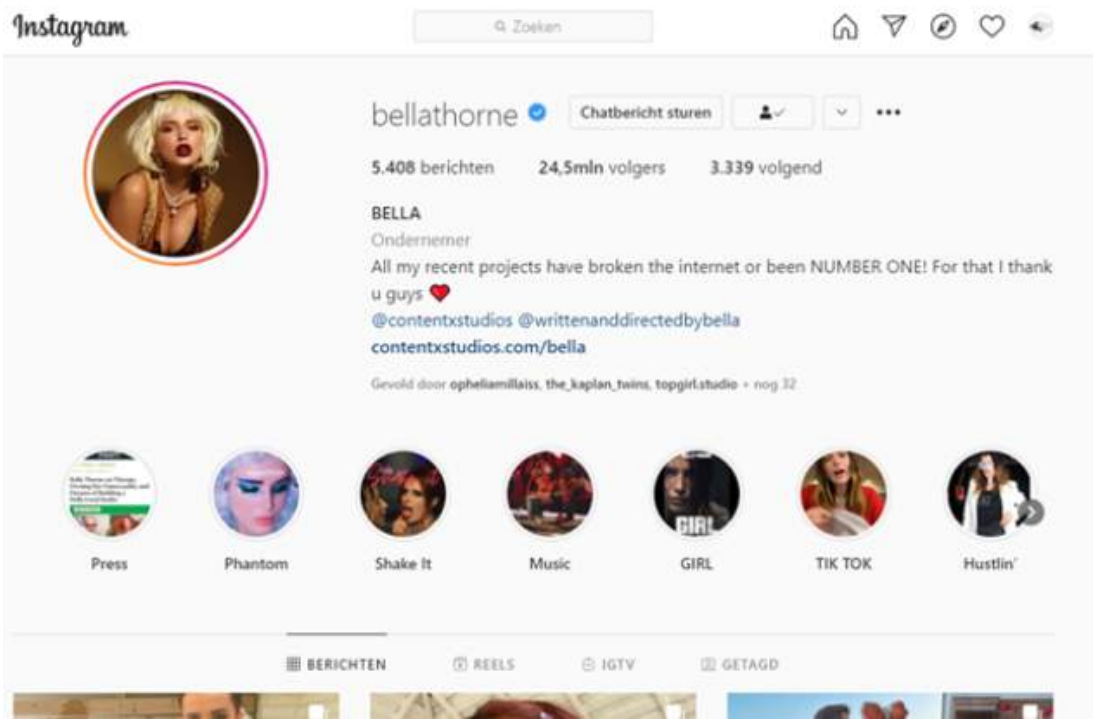

Fig 2. Philippo, Lisa. "Screenshot of the *Instagram* Webpage of Bella Thorne", *Instagram*, May 11 2021, `www.instagram.com/bellathorne`, screenshot.

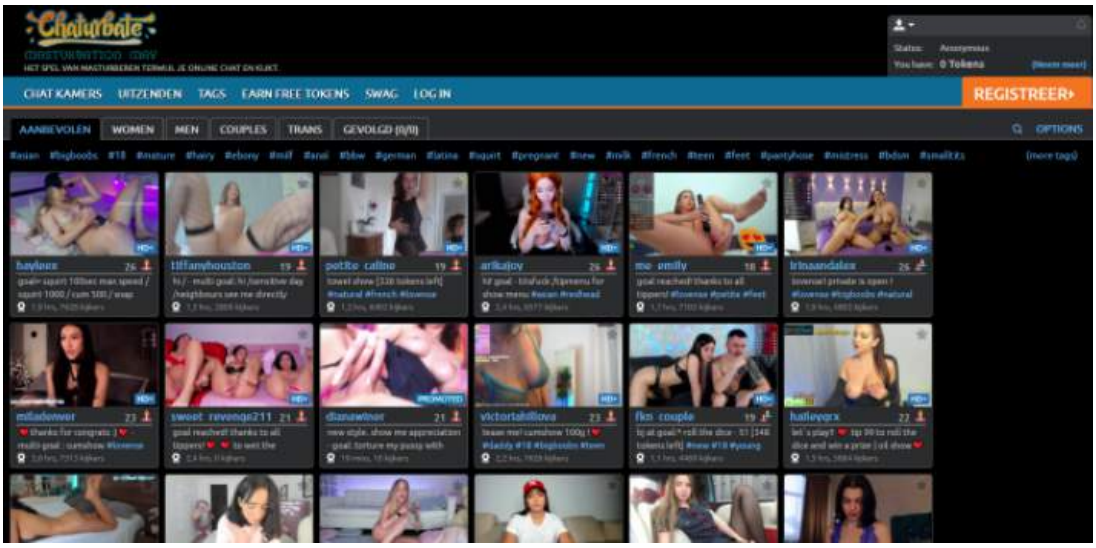

Fig 3. Philippo, Lisa. "Opening Interface of *Chaturbate*", May 11 2021, `www.chaturbate.com`, screenshot.

Humanities

Remembering Soviet Lithuania in Grūtas Park

A Case for Counter-monumentality

Miglė Gerčaitė

Supervisor
Prof. Dr. Dawn Skorczewski (AUC)
Reader
Dr. Marco de Waard (AUC)

Photographers: Daria Roman and Richard Essink

Abstract

Grūtas Park represents a unique location in the memory landscape of Lithuania. Housing a significant material heritage of the Soviet period, the park exhibits and parodies monuments that were removed from public spaces in Lithuania after the country achieved its independence from the Soviet Union (USSR) in the 1990s. The controversial reception of Grūtas Park in Lithuanian society, however, offers clues to its contributions to the complex and multidirectional memory of Soviet Lithuania and the park's insufficient commitment to counter-monumental memory work may contribute to public disagreement over its memory representations of the Soviet occupation. To what extent are the park's subject, form, visitor experience, site, and meaning counter-monumental? Do its exhibits create a multifaceted narrative of Soviet Lithuanian history, conveying it to visitors through diverse sensory experiences? Introducing the dialogic approach as a way to interrogate the role of counter-monumentality in Grūtas Park, I propose interaction between memory narratives at Grūtas Park and public spaces in Lithuania as a means of working through conflicts over memory concerning the Soviet occupation. I argue that Grūtas Park engages with both traditional and counter-monumental forms of representation, encapsulating nationalistic and patriotic historical narratives, trauma of the Soviet period, nostalgia for daily life in Soviet Lithuania, and ironic and carnivalesque portrayals of the Soviet system. Having established the foundations for a counter-monumental representation of Soviet history in Lithuania, Grūtas Park may ultimately offer a comprehensive and esteemed representation of Soviet memory in Lithuania.

Keywords: *Anti-Nostalgia, Counter-monumentality, Cultural Memory, Lithuania, Soviet Occupation*

Suggested citation: Gerčaitė, Miglė (2021). "*Remembering Soviet Lithuania in Grūtas Park: A Case for Counter-monumentality*". AUC Undergraduate Journal of Liberal Arts and Sciences, vol. 16, no.1, pp. 117-136

Contents

List of Figures and Tables

1 Introduction

In the southern region of Lithuania, near the city of Druskininkai, stands a big park named Grūtas Park (see fig. 1). Hosting an outdoor exhibition of statues and other memorabilia from the period of Soviet occupation, the park offers a unique location in the memory landscape of Lithuania. It accommodates numerous Soviet monuments that were removed from public spaces across Lithuania in the 1990s, after the country gained independence from the Soviet Union (USSR). Entering the park, visitors see a replica of the freight-train cars used in the 1940s to transport citizens to Siberia and other regions to work and live under harsh conditions (see fig. 2). Visitors also encounter a number of watchtowers and barbed wire that represent the imprisonment of individuals and the nation as a whole. Yet, the park also resembles a theme park: speakers play Soviet marching songs, a café offers 'authentic' Soviet dishes, and a kiosk sells a newspaper *Grūto parko tiesa (The truth of Grūtas Park)* – a take on the Soviet newspaper *Pravda (Truth)*, which disseminated propaganda (*Grūto Parkas*). The monuments at Grūtas Park memorialize the painful history of Soviet occupation, but they also recontextualize it, stripping the monuments of their ideological and political weight through parody and caricature.

Figure 1: Entrance to Grūtas Park. Žygimantas Gedvilas, 15min.

Deliberately opened on April Fools' Day 2001, the parody of reverence for the Soviet past in Grūtas Park represents important elements of the Soviet occupation of Lithuania. Annually, on May 9th, the park stages theatre-like parodies of Soviet figures such as Stalin, Lenin, and others (Williams 189). Exhibits at the park memorialize a difficult history of occupation and imprisonment, but also invoke a sense of nostalgia for daily life in Soviet Lithuania. A Soviet era playground for children, *Luna Park*, invites youth to imagine how their parents used to play in Soviet Lithuania. A brightly lit café offers the opportunity to try out their 'Soviet menu' and experience 'authentic' Soviet cuisine (*Grūto Parkas*). Since it reaches in two directions, towards pain and nostalgia, the park has received mixed receptions from visitors. Some view the light-hearted aspects of the park as tone-deaf and disrespectful to victims of the Soviet regime, while others appreciate the more cheerful Soviet nostalgia (Isaac and Budryte 7).

Figure 2: Replica of a freight-train car. Žygimantas Gedvilas, 15min.

Nonetheless, Grūtas Park offers a solution to the issue of remembrance of Soviet history, which arose in the 1990s with Lithuania's independence from the USSR. However, it also poses serious questions about the nature of memorialization itself. This analysis of the Grūtas Park case questions what the park achieves in terms of memory work and memory discourse in Lithuania. It debates whether the park may provide insights into current memory debates and serves as an example of the applicability of the counter-monumental approach to problems of cultural memory. Grūtas Park constitutes an important case for memory studies and analysis, for it attempts to represent a complex and controversial era of Lithuanian history. James E. Young's theory of counter-monumentality provides an apt model to analyse how Grūtas Park conducts memory work (i.e. engages with the past), as well as the broader significance of this engagement in discursive practices in Lithuania.

Differently from the traditional monument, which affirms reality and conducts memory work on behalf of the viewer, the counter-monument provides space for changing representations and continuous memory work done by the viewer themselves (Young, "At Memory's Edge" 372, 374). The counter-monument does not affirm specific memory narratives, but instead explores notions of doubt and impermanence in historical remembrance ("At Memory's Edge" 374). Specifying the theory of counter-monumentality, Quentin Stevens et al. suggest a distinction between 'anti-monumental' and 'dialogic' forms of counter-monumentality (719). Adopting this distinction in its analysis, this paper aims to examine to what extent Grūtas Park exemplifies notions of counter-monumentality both in anti-monumental and dialogic forms. It questions what kind of memory work is encouraged at the park and whether this creates a congenial space for contrasting memory narratives.

1.1 The Historical and Geographical Context of Lithuania

As a small country by the Baltic Sea, with powerful neighbours such as Russia and Poland, Lithuania is seldom recognized in historical narratives of the 20th century. The

country's tumultuous and intriguing history is marked by continuous attempts at nation-building, grappling with occupation, and an unrelenting fight for independence. First declaring independence from the Russian Empire in 1918, Lithuania emerged as a modernizing nation-state during the interwar period (Davoliūtė 2). Lithuania's territorial integrity was first challenged in March of 1939, when Nazi Germany gained control over the region of Klaipėda. In August of the same year, Nazi Germany and the Soviet Union signed the Molotov-Ribbentrop Pact, which delineated Soviet and German spheres of influence in Eastern Europe. Hence, Lithuania was briefly occupied by the Soviets in 1940, experiencing mass deportations, but as war broke out between the Soviet Union and Nazi Germany, Lithuania came under Nazi control in 1941. Following the genocide of Lithuanian Jews, mass killings, and emigration, the Soviet Union again occupied Lithuania in 1944. The country remained occupied by the USSR for almost 50 years. Thus, the 1940s were a significant period of Soviet Lithuanian history, when many Lithuanian citizens were deported to remote locations, most notably Siberia. People of different demographics lived and worked under harsh conditions, while political dissidents were imprisoned in labour camps, known as Gulags. The history of deportations later became an important feature of the independence movement of the 1980s, which ultimately led to the re-establishment of the Lithuanian state in 1990 (Davoliūtė 155).

During the Soviet occupation, the artistic and cultural landscape of Lithuania underwent many changes. Aiming to erase Lithuanian national identity and replace it with a Soviet one, the Soviet state built monuments and cultural centres conveying communist ideology, and introduced new holidays such as the anniversary of the October Revolution (Černiauskaitė and Jasiūnaitė). Furthermore, the national anthem of Lithuania was banned, and replaced with a Soviet version that praised Lenin, communism and the communist party (Antanavičius). At the same time, opposition to Sovietization and campaigns for independence assumed political and artistic forms. In 1972 a young student Romas Kalanta publicly self-immolated in Liberty Avenue, Kaunas in protest against Soviet rule (Davoliūtė 132). Following this, public demonstrations and protests broke out, which later became known as Kaunas Spring. Anti-Soviet Catholic activists produced an underground publication – *The Chronicle of the Catholic Church in Lithuania*, which reported on the oppression of Catholics in Soviet Lithuania, as well as other crimes against the people by the Soviet state (*Lietuvos Katalikų Bažnyčios Kronika*). *Sąjūdis*, a political organisation formed to fight for independence, published numerous dissident newspapers, founded a commission for historical investigations of crimes against the Lithuanian people under Joseph Stalin, and sponsored historians to write a Lithuanian national history (Makhotina 3).

1.2 Anti-Soviet National Identity and Cultural Memory

The history of Soviet Lithuania spans almost 50 years. Many historical events from the Soviet era are considered part of the collective memory of Lithuanians, and are instrumental in delineating national identity. Ernest Renan defines the nation as 'a soul, a spiritual principle' (19). This means that the nation is classified not by its laws or territorial claims, but rather the will and beliefs of the people. The people must share a 'common legacy of memories,' as well as consent to live together and consider themselves a nation (19). For this reason, Renan interprets the nation as a 'daily plebiscite' (19). It is important to recognize that the history and memory of a nation is subjective, and hence, the essence of a nation is that its people "have many things in common, and also that they have forgotten many things" (11). In Lithuania, national identity is outlined by remembrance of specific historical events, which depict the suffering and perseverance of Lithuanian people. The trauma of deportations, the 1944-1953 partisan warfare, acts of dissent by people such as Romas Kalanta and organisations like *Sąjūdis*, are all remembered and considered part of a Lithuanian national identity. At the same time, unpleasant or shameful memories such as the participation of Lithuanians in the Holocaust, the support for Soviet ideologies by Lithuanian writers and politicians, as well as modernization and economic development due to Soviet policies are all forgotten or repressed.

The Soviet occupation remains a defining aspect of remembrance practices in Lithuania today, in part remembering a painful past, and in part serving to reaffirm Lithuanian national identity. For instance, June 14th is known as Gedulo ir Vilties Diena (Day of Hope and Mourning) and commemorates the 1940s deportations. It is usually memorialized by the *Ištark, Išgirsk, Išsaugok* (*Say, Hear, Save*) action, where live readings of deportee names are held in various locations across Lithuania (*Ryžtas žydėti*). In addition, the Mission Siberia project receives widespread public support and media attention every year. The project founders organise annual expeditions to locations in Russia, Kazakhstan, where deportees were buried in order to maintain Lithuanian cemeteries (*Misija Sibiras*). These memorial activities serve as remembrance of Soviet crimes and the lives lost. Yet, they also represent a contemporary cultural solidification of Lithuanian identity, in which remembering suffering from the past brings people together.

Importantly, such remembrance practices are framed within discourses of patriotism and nationalism. The 1980s independence movement, led by *Sąjūdis*, was characterized by calls to reclaim Lithuanian national identity and re-establish the Lithuanian nation state. Nationalism developed in opposition to communism, and strongly influenced the history and memory culture of Lithuania (Michnik and Matynia 759). Experiences of deportations were understood as identifiers of belonging to the Lithuanian nation, and served as a basis

for political mobilisation. Consequently, the commemoration practices in modern-day Lithuania convey ideas of patriotism, with *Mission Siberia* stating that their goal is to "build social awareness and patriotism" (*Misija Sibiras*). Significantly, the validity of such a mindset is in part legitimized by the geopolitical situation. Political tensions with Russia remain to this day, borne of divergent interpretations of key moments in Soviet history. Whereas the 1940s are viewed as the period of victory against Nazism in the Great Patriotic War in Russia, Lithuanians grimly remember it as the beginning of Soviet occupation and the replacement of one authoritarian regime with another (Marciniak 15). On May 9th, when Russian-speaking communities in Lithuania celebrate Victory Day and hold demonstrations, Lithuanian nationalists often respond with counter-demonstrations, leading to conflict and extensive media coverage (Grigaliūnaitė, "Gegužės 9-ąją").

1.3 Unwanted Heritage: Soviet Monuments in Independent Lithuania

In the memory landscape of Lithuania, frequent and deliberate memorialization of the Soviet past is continuously taking place. Within this context, Soviet monuments that were built during the Soviet occupation and remained as legacies of the regime, occupy a difficult position. The Soviets created many ideological spaces and objects, such as buildings, statues, paintings, books, etc. For example, in the capital of Lithuania, Vilnius, two main squares were remodelled for ideological purposes. In 1952, Lukiškės Square was renamed to Lenin Square and a large statue of Vladimir Lenin rose up at its centre. A monument for the writer Elzė Ožeškienė in Vincas Kudirka Square was demolished, and replaced with a memorial to a Soviet Army general Ivan Chernyakhovsky (Grigoravičienė). Similar remodelling took place in other major cities such as Kaunas or Druskininkai as well, where statues of Lenin appeared in main city squares (Grigoravičienė). However, in the early 1980s and 1990s, the Lithuanian independence movement began targeting Soviet monuments in attempts to reclaim cultural memory. *Sąjūdis*, which was looked to as the ideological authority of the Lithuanian independence movement, pledged to remove all Soviet statues and monuments from public spaces (Makhotina 6). Doing so would demonstrate a rejection of communist ideology, as well as erase Russian presence in Lithuania (6).

Large jubilant crowds celebrated in Lukiškės Square in Vilnius as the Lenin statue toppled to the ground, and similar monumental celebrations took place across Lithuania (Makhotina 6). In the early 1990s, many statues were lost or destroyed, while those that remained were stored in various warehouses and courtyards across the country. However, by the late 1990s, the government of a now independent Lithuanian state began deliberating over how to re-incorporate Soviet monuments into the Lithuanian memory landscape. The state feared that without monuments to serve as evidence of the Soviet past, history would be erased and forgotten (Lankauskas 29). Yet, the government could not decide how Soviet history should be remembered and preserved (Kuczyńska-Zonik 146). Eventually, the Lithuanian state chose to leave the decision in the hands of private stakeholders by hosting a contest for projects that would address the monument problem (Williams 187). Viliumas Malinauskas won the contest with his project of Grūtas Park. Arguably, the most appealing aspect of the project was the fact that it would be privately funded by Malinauskas himself, since the state believed that taxpayers would be unwilling to dedicate funds to preserve symbols of the Soviet regime (187). The project of Grūtas Park was approved in 1998, and it now houses 86 Soviet monuments, as well as other Soviet memorabilia (*Grūto Parkas*). Considering this context, in order to examine Grūtas Park and its memory discourses, it is first necessary to establish an analytical framework based on memory studies.

2 Research Context

This paper applies James E. Young's theory of counter-monumentality, in which he defines the counter-monument in opposition to the traditional monument, to the case of Grūtas Park. For Young, the monument reflects the socio-historical and aesthetic context, and is responsive to the needs of both art and official history (Young, "At Memory's Edge" 372). In the 20th century, the traditional monument became popular due to its appeal to authoritarian regimes, since it acted as a reaffirmation of specific ideologies ("At Memory's Edge" 373). Such monuments have an 'essential stiffness,' and represent a clearly delineated, fixed meaning. They possess designated authority over history and memory, affirming permanent reality and essentially doing the memory work for viewers ("At Memory's Edge" 372). The counter-monument, on the other hand, rejects notions of authority over memory and has the capacity to represent changing and undefined memory ("At Memory's Edge" 373). With such monuments, the burden of memory work falls onto the viewer, as the monument itself generally does not represent a specific ideology ("At Memory's Edge" 374). To this end, the counter-monument may challenge reality instead of affirming it, exploring notions of doubt and impermanence in historical remembrance ("At Memory's Edge" 374). To analyse the ways in which Grūtas Park encourages counter-monumental memory work and accommodates differing interpretations of memory, it is productive to consider the 'anti-monumental' and 'dialogic' distinction of counter-monumentality.

Stevens et al. argue that various terms such as 'non-monument' or 'deconstructive' have been used to refer to the notion of counter-monumentality with additional meanings and connotations (719). Therefore, they suggest that a distinction between the anti-monumental and the dialogic is useful to more accurately analyse the way counter-monuments function (718). Anti-monumental forms utilize unconven-

tional techniques of monumentality, representing subjects and ideas that are not typically addressed by traditional monuments (719). The image of a 'pathetic hero' is an example of an atypical, counter-monumental subject in the 20th century context, since during that time, authoritarian regimes solely focused on producing traditional monuments representing valorising images of war victories (Young, "At Memory's Edge" 372). The dialogic denotes a more direct relationship with existing monuments, acting as a critique of a specific monument and its meanings (Stevens et al. 719). A counter-monument is "intentionally juxtaposed to another, pre-existing monument located nearby," and its represented meaning addresses and critiques "the values the preexisting monument expresses" (729). The dialogic, then, is concerned with new meanings that the interaction between two monuments produces (729). In order to identify counter-monuments and distinguish them from traditional monuments, Stevens et al. suggest specific dimensions of subject, form, site, visitor experience, and meaning (722). This paper will analyse the ways in which exhibits and other attractions at Grūtas Park engage with these dimensions and the dialogic form to indicate how counter-monumentality is facilitated at the park. However, it is important to first clarify the complexity of Soviet memory in Lithuania to illustrate how this influences the effectiveness and reception of counter-monumental representations at the park.

This paper conceptualizes Soviet memory as multidirectional, which offers conflicting interpretations and evokes intense feelings for different stakeholders. In his conceptualization of multidirectional memory, Michael Rothberg aims to address the perception of memory as competitive, which legitimises tensions caused by conflicting historical narratives in public discourses (2). In public spheres, different collective memories are compared and contrasted in a 'zero-sum struggle,' meaning that specific memory narratives are perceived as the only acceptable ones (3). Rothberg concedes that in certain instances, it is difficult to reconcile differing memories and prevent deterioration into violence (3). However, this does not imply that memory is fixed and competitive. It is multidirectional, meaning that it is "subject to ongoing negotiation, cross-referencing, and borrowing" (3). Memories are constantly moving and changing, and interacting with each other in productive ways. For Grūtas Park, this suggests that referential memory work related to Soviet history is possible at the park, and may lead to productive results.

However, in contrast to Rothberg's conception, Soviet memory in Lithuania escalates tensions in society (Matonytė 7). Two main 'patterns of discourse' by which people cope with and relate to the Soviet past – 'anti-nostalgia' and 'nostalgia' are opposed to one another in public debates (108). This is significant, as Grūtas Park attempts to represent both anti-nostalgic and nostalgic sentiments in its exhibition. This essay questions whether the tensions in society between these two patterns of discourse are exemplified at the park, and inform decisions of representation, and consequently, engagement with counter-monumentality.

Anti-nostalgia is defined by a renunciation of the Soviet past and a complete rejection of Soviet ideology (Matonytė 108). Public discourse in Lithuania is largely characterized by the anti-nostalgia framework, meaning that people view the entire Soviet period as negative, and believe that the current social and cultural conditions of Lithuania are preferable. Irmina Matonytė suggests that anti-nostalgia manifests in cultural and public policies of the independent Lithuanian state. She notes that in the 1990s the Lithuanian parliament *Seimas* passed a number of laws prohibiting former members of Soviet secret services from participating in public service or education (Matonytė 108). *The Centre for Genocide and Resistance Studies* held various conferences and made publications informing the public of the trauma of the Soviet occupation (109). The Centre systematically documented narratives of victimization, encouraging a complete rejection of the Soviet period.

Therefore, the anti-nostalgic discourse must to an extent be represented at Grūtas Park, since it exemplifies the official position of the original facilitator of the project, the Lithuanian government. The current government is seemingly continuing the 1980s and 1990s politics of *Sąjūdis*. During this period, the history of 1940s and 1950s deportations became acutely relevant, as deportee memorials were published and historians wrote new publications of Lithuanian history (Davoliūtė 168). The trauma of deportations empowered a complete rejection of the entire Soviet period, providing no space for consideration of positive developments under the regime (159). *Sąjūdis* initiated commemorative practices of deportation history, which are continued by the Lithuanian government, such as the Mourning and Hope Day on June 14th that commemorates the first deportations. This, and various other memorials consolidate anti-nostalgic attitudes in society, while also indicating the presence of nationalist ideologies (Matonytė 108).

On the other hand, the nostalgic discourse stems from personal experiences and autobiographical memories of Lithuanian citizens, providing a more positive interpretation of the Soviet period (Matonytė 111). It largely appears among smaller social groups of Lithuanian citizens, characterized as fringe groups by the state. It is expressed through an appreciation of certain aspects of the Soviet era, or may even manifest as an 'uncritical glorification' of the period (111). A number of Lithuanian people were able to live relatively 'normal lives,' which were not obstructed by ideological oppression of the Soviet state or suppression due to political activism (Klumbytė, "Memory" 304). Reflecting on the past, Lithuanian citizens identify work opportunities, secure financial income and collective self-identification as positive aspects of the Soviet era, now lost in the highly individualized, economically uncertain post-Soviet present ("Memory" 304). At Grūtas Park, nostalgia is embraced by portraying childhood memories at *Luna Park* or offering a Soviet gas-

tronomic experience at the café. Even though memories of 'good Soviet times' do not necessarily relate to political or ideological interpretations of the Soviet and post-Soviet state organisation, they are nevertheless 'misrecognized and stigmatized' on both state and societal levels ("Memory" 295). For instance, in 2009 the Lithuanian government passed a law, which prohibited "propagating, denying, or harshly depreciating or approving communist or fascist genocide as well as other crimes," the violation of which could lead to imprisonment for up to three years ("Memory" 296). According to Klumbytė, the state may misrecognize memories of 'good Soviet times' as a denial of 'communist atrocities,' and hence, criminalize them ("Memory" 296). A number of politicians and intellectuals believe that "those who remember Soviet times vote populist and think communist" ("Memory" 296). Therefore, citizens holding positive memories of Soviet times are stigmatized and even considered a threat to democracy and the state's independence ("Memory" 296).

The state and media view Soviet nostalgia as an indication of bad citizenship, which results in a marginalisation of a part of Lithuanian society that may hold such views (Klumbytė, "Memory" 307). For instance, Klumbytė carried out interviews in village communities in Lithuania, where she encountered the phrase "Prie ruso buvo geriau" ("It was better with the Russian") (Klumbytė, "Dabarties istorijos" 5). She found that the phrase was a 'social idiom,' a popular statement, recognizable through common knowledge in village communities and expressing a positive outlook on the Soviet past ("Dabarties istorijos" 18). Klumbytė states that, while commenting on her interview findings, some Lithuanian politicians had stated that the positive Soviet reality encapsulated by "Prie ruso buvo geriau" is "not the reality which political parties are striving for" ("Dabarties istorijos" 20). In mainstream cultural and political spheres, the phrase has indeed become widely recognizable as the epitome of nostalgic sentiments and is often referred to in a mocking way to criticize the nostalgic attitude (*"Prie ruso buvo geriau"*). The phrase, tinged with irony, can be considered a cultural adage, suggesting that believing in such things is simply incomprehensible. In its attempts to reach a sense of closure with the Soviet period and consolidate independence, the Lithuanian state generates antagonism between the anti-nostalgic and nostalgic narratives, resulting in further societal polarization and marginalization (Gussi 126). Rather than engaging in honest discussions and recognizing the subjective nature of nostalgic sentiments, media outlets and government officials outwardly reject any positive associations, and use tactics such as the "Prie ruso buvo geriau" adage to categorize all positive feelings as irrational and unacceptable.

Therefore, the confrontation between anti-nostalgia and nostalgia may also be considered a conflict between institutional and individual memory narratives, applying Jan Assmann's concept of cultural memory and communicative memory as two *modi memorandi* (Erll 30). It is important to identify the different ways that cultural and communicative memory function in Lithuanian society, as Grūtas Park engages with both. As a government sanctioned, educational institution, the park aims to live up to governmental expectations, representing institutionally supported memory narratives. At the same time, it is a private entity, which seeks to attract visitors by appealing to their personal sentiments in order to continue making a profit.

Jan Assmann identifies criteria that constitute cultural memory, which are often visible in the practices of commemoration by official institutions. Cultural memory carries a sense of 'obligation' that necessitates its preservation and continuous remembrance (J. Assmann 214). This burden of conservation falls on authorities such as teachers and historians, as well as institutions such as museums or schools (212). Consequently, collective identity is strongly linked to cultural memory by a 'concretion of identity' (213). Institutions and people identify certain historical events as part of a national history, and their remembrance becomes a requirement for national identification. Moreover, cultural memory has a 'capacity to reconstruct,' meaning that it not only reminds a people of certain events or memories, but it also contextualizes them in a specific way, for example, mourning victims of a war or celebrating war victories (J. Assmann 214). Finally, cultural memory has a capacity for 'reflexivity,' meaning that it reflects the cultural group's "lifeworld and its self-image, and is moreover self-reflexive" (Erll 30). Cultural memory is 'practice-reflexive' by illustrating common practices through the use of idioms and proverbs, 'self-reflexive' as it 'draws on itself' to interpret and critique, and 'reflexive of its own image,' as it "reflects the self-image of the group through a preoccupation with its own social system" (J. Assmann 215). In Lithuania, the historical narrative of deportations and anti-nostalgic attitudes function as cultural memory. As an authority on history, the Lithuanian government considers this memory as part of a national history, linking it to national identity by establishing commemorative practices on an institutional level. Commemorating the history of deportations by taking part in memorial events on days such as the Hope and Mourning Day, may qualify as a requirement to identify as a Lithuanian.

In contrast, communicative memory is more individualized and accommodates personal experiences and memories within a social community (J. Assmann 212). Communicative memories are biographical, and hence, have a limited temporal scope (Erll 28). Interviewing people about their lives during the Soviet times, Neringa Klumbytė showcases how positive memories are linked to specific personal experiences, which indicates their communicative nature ("Memory" 298). Having a working Soviet car, savings to buy clothes, visit a resort during holidays or buy a private house, are all positive aspects that people conflate with the Soviet period ("Memory" 298, 300, 301). Since people believe that such opportunities are no longer possible in the current economic climate, they express feelings of nostalgia in regard to the past, and wish for better times. Intrinsically, there is

no opposition or overlap between the cultural memory of deportations and the communicative memory of daily life. For Jan Assmann, the communicative and cultural dimensions of memory do not directly oppose one another (Erll 30). However, the Lithuanian state's cultural events and legal policies convey its antagonistic approach towards any positive aspects of the Soviet times.

As mentioned before, in 2009 the Lithuanian state passed a law, which prohibits denial or appreciation of communist crimes. Klumbytė argues that as a consequence, the state is able to argue that memories of 'good Soviet times' deny the atrocities of the Soviet regime, and hence, may be criminalized (Klumbytė, "Memory" 296). Therefore, the state may actively incite conflict by criminalizing nostalgic aspects of communicative memory that do not conform to the cultural memory's narrative of anti-nostalgia. It is necessary to recognize this friction in society, as it informs both the ways Grūtas Park presents its exhibition and how it is perceived by the public. Discourses of Soviet atrocities and positive memories are both present within society, further suggesting that Soviet memory is indeed complex and multidirectional. Grūtas Park is tasked with the responsibility of incorporating these representations in its exhibition and, by attempting to do so, it delves into counter-monumental forms of representation. Thus, in order to identify the potential ways in which Grūtas Park explores counter-monumentality, it is productive to consider analyses of the park by previous researchers.

Ekaterina Makhotina suggests that as a memory site, Grūtas Park embraces both alienation and nostalgia (7). The ironic, humorous approach to representing history through parodies of Soviet leaders serves to alienate the viewer from the Soviet past, making Soviet ideology inconceivable through its silliness. Other aspects of the park, such as the café, evoke feelings of nostalgia (7). For Katarzyna Marciniak, this ironic contextualization of monuments at the park creates a sense of otherness, defining the park as a heterotopic space, a cemetery of 'exiled rejects of history' (16). She argues that the Soviet period is 'othered' and dissociated from modern day Lithuanian society and history. In contrast, Gediminas Lankauskas determines that instead of a cemetery, Grūtas Park functions as a place of imprisonment, where the visitor assumes a position of dominance and authority over the 'imprisoned' monuments (38). Whereas the term 'cemetery' refers to an 'execution,' an end to the Soviet regime, the term 'imprisonment' implies a defeat, and an authoritative position of visitors as representatives of the current political system (38).

From another perspective, Paul Williams is sceptical of Grūtas Park's engagement with memory work and discursive narratives (194). For him, the park works as a place of 'banishment,' where monuments are removed from the political and geographical landscape of Lithuania (194). Removing monuments from public space represents a definite end to the Soviet era, as well as a rejection of its ideology, yet, it does not ensure a continuous discussion of its legacy (195). Like Williams, Alexandra Kuczyńska-Zonik argues that Grūtas Park provides a limited space for memory work, since the monuments have been 'neutralized' by their removal from public spaces, and stripped of any authority or prestige (155). Overall, there seem to be spheres of contention in understanding the park's memory work. The effect of parody and irony on the viewer and remembrance, the significance of the park's site as a peripheral space, and the capacity for memory work and the viewer's relationship to the monuments are all aspects, which will be analysed within the framework of counter-monumentality.

3 The Controversial Establishment of Grūtas Park

The project idea for Grūtas Park was controversial from the very beginning. The development and foundation of the park were followed by critical societal debates. Some people viewed the project as a beneficial and important attempt at preserving history, while others deemed it profane and even criminal (Lankauskas 29). It was argued that building such a park would not only bring back "haunting memories of one of the most horrifying periods of Lithuanian history," but it would also be an insult to the memory of those who had suffered under Communist rule (Isaac and Budryte 7). An example of the criticism during the planning of the park is the frequent questioning of the park's location. The park was to be built near Grūtas Forest, which was used as a base by a number of Lithuanian partisans during the 1945-1953 partisan resistance movement (7). The placement of Soviet monuments and the exhibition of the Soviet regime in such a historically significant location was viewed as disgraceful and offensive (7). Moreover, there were concerns that the region of Dzūkija, where the park would reside, would become a gathering spot for former communists and sympathizers of the Soviet regime ("2001").

Opposition was also visible at the governmental level. Some members of parliament, led by Juozas Galdikas, worked towards having the park banned and shut down (Dapkus). Galdikas criticized the founder of the park, Malinauskas, arguing that both the location and concept of the park were unacceptable: "Malinauskas does not care that these forests where Grūtas Park was built once served as a shelter for Lithuanian freedom fighters against Soviet occupants. He does not care about the painful history of Lithuania" (Williams 193). Furthermore, the comedic and parodic approach of the park's exhibition was not at all amusing for Galdikas or his supporters, who asked: "What is the purpose of this park? To laugh at our pain?" (193).

Thus, the founders and curators of the park were forced to convince the public and lawmakers that the purpose of Grūtas Park was not to glorify the Soviet past and its leaders,

but rather to expose the harsh truth of Soviet ideology and serve as a reminder of history ("2001"). Various Lithuanian communities suggested that the park could provide 'valuable historical lessons' for future generations (Isaac and Budryte 7). In response to criticisms by Galdikas and others, Malinauskas stated that due to his family history of deportations he understands "all the fun that Uncle Stalin and his cronies provided for us," adding that the park commemorates an important part of Lithuanian history in a unique and engaging way ("2001"). In his opinion, "If this park wasn't interesting, people wouldn't come here," and hence, they would not be exposed to Soviet history ("Stalin World"). Moreover, the caricatures of Soviet leaders and symbols, and the ability to laugh at them signifies that Lithuanians are no longer intimidated and controlled by the communist regime (Dapkus). Malinauskas even erected wooden statues of parliament members that petitioned against the park as part of Grūtas Park's exhibition, stating that "those who are still afraid of shadows of the past deserve to stand here" (Williams 193).

Attempts at curtailing the project eventually failed, and the park officially opened in 2001 amidst protests by radical patriots ("Melagių dieną"). After its opening, public reception of the park's exhibition was also mixed. At Grūtas Park, Rami K. Isaac and Laurencija Budryte analysed visitor comment books, i.e. books usually placed in exhibition halls, where visitors can write down their thoughts and feelings about the exhibition, as well as read other visitors' messages. The researchers indicate the often contradictory feelings and opinions expressed by visitors of the park. Some visitors felt enthusiastic about the park and the exhibition, valuing its creativity (Isaac and Budryte 10). They were impressed by the fact that 'unwanted history' was being represented, and believed that the park would memorialize history for future generations (17). Others were unhappy with the park's execution, for instance, criticizing the fact that important historical figures such as the leaders of the Lithuanian Communist Party, Antanas Sniečkus and Algirdas Brazauskas, were omitted from the exhibition (13). For these visitors, the park had a responsibility to enact historical justice, and failed to do so. Others still felt discomfort and even emotional shock due to the horrible parts of history that the exhibits memorialized (20).

Some survivors of deportations expressed strong feelings of anger and disgust at both the park and its creators. Povilas Dirsė, who had protested against the establishment of the park, stated that the park was not needed, as books and other educational sources could provide explanations of Soviet history ("Stalin World"). He believes that the park is a harmful, commercial space: "When tours go there, that's propaganda, pro-Soviet propaganda. I'm against it" ("Stalin World"). Another deportee, Antanas Terleckas, does not see the value in preserving Soviet monuments at all: "These are symbols of the beastly Communist regime. I mean of [*sic*] a political organization that engaged in genocide. So why keep them? ("Stalin World"). Asked whether she had visited the park, Nijolė Sadūnaitė replies: "I haven't been there and I'm not going to" ("Stalin World"). Prominent political figure and former chairman of *Sąjūdis* Vytautas V. Landsbergis expressed a similar, emotionally charged opinion about the park: "Why? For what reason do we have to keep those symbols and idols of a stupid and criminal ideology, and our conquerors and executioners? This is crazy." ("Stalin World").

Intellectuals and representatives of academic circles are also doubtful of the park's contribution to the remembrance of Soviet history. Commenting on the creative vision of the park, art critic Laima Kreivytė suggests that Grūtas Park is essentially a single person's vision of how to deal with Soviet memory and heritage ("Parko Gimimas ir Mirtis"). Hence, she questions the extent to which the park can provide a space for visitors to engage with history on their own terms. For Kreivytė, the park's representation of history is too static, limiting the capacity for future generations to re-define remembrance and their personal relationship to Soviet history and the park. Cultural historian Salvijus Kulevičius also believes that the founder Malinauskas had a significant role in the creative vision of Grūtas Park. Therefore, he is doubtful of the park's engagement with notions of openness and publicness ("Parko Gimimas ir Mirtis"). Simply put, the park remains an undecided and controversial space, which prompts differing, and often, very strong emotional responses.

4 Anti-monumental Dimensions at Grūtas Park

Analysing Grūtas Park, I argue that the park exemplifies anti-monumental forms of memory representation. Through the anti-monumental dimensions of subject, visitor experience, and form, the park conveys multidirectional memory narratives, which entail counter-monumental memory work. The status of Soviet history as an unresolved and controversial memory, the parodic and carnivalesque visitor experience, as well as monumental forms that subvert power relations, indicate a counter-monumental direction in representation. At the same time, I suggest that memory work at the park becomes limited and measured, because the park engages with traditional monumental forms. Site and meaning illustrate a want to conform to governmental expectations, such as the anti-nostalgic framework. A peripheral location creates distance from public discourse, and the focus on representing a preferred memory narrative demonstrates the limits of counter-monumental memory work at the park.

4.1 Subject: Painful History and 'Hot' Discourse

Stevens et al. suggest that in contrast to traditional monuments, anti-monuments often represent atypical sub-

jects in history and memory. Since traditional monuments are affirmative and glorifying, they usually celebrate positive achievements in history such as war victories (Stevens et al. 722). Traditional monuments typically commemorate exceptional historical figures and convey a narrative of heroism and bravery (722). Anti-monuments, on the other hand, most often reflect negative events and tragedies, or raise doubts about certain historical narratives and truths (722). In representations of war, anti-monuments concentrate on commemorating the victims or condemning the perpetrators of crimes (722). The Holocaust is an exemplary subject of anti-monumental work, since Young's theory of counter-monumentality was, to a large extent, developed by reflecting on Holocaust memorials in Germany and their focus on the suffering of victims, self-reflection of the perpetrator nation, and continuous memory work (Young, "The Counter-Monument" 267). The exhibition begins by telling the story of human suffering and injustice. The first monumental object that the visitor sees is a freight-train car that replicates cars in which Lithuanian citizens were taken to Siberia in the 1940s (Lankauskas 38). A plaque on the car explains that the object represents a time when the Soviet state was "carrying out massive genocide of the Lithuanian nation" (39). The exhibit calls for contemplation and solemnity, illuminating the tragic side of 1940s history. Moreover, the space is surrounded by a few watchtowers and is fenced off by barbed wire, creating a sense of a "penitentiary panopticon" (38) (see fig. 3). Venturing further into the park, the visitor sees a few monuments placed inside a field that is encircled by barbed wire. The monuments are inaccessible to visitors for closer inspection, as if they themselves are imprisoned. In this way, the park simultaneously conveys a sense of entrapment, recalling perhaps the most painful part of Soviet history, as well as physically and metaphorically ensnares the Soviet monuments, presenting them as objects of the past, now inaccessible to the modern viewer. Such staging of exhibits creates a sombre atmosphere and invites visitors to stop and contemplate the history. I suggest that the subject of the park is anti-monumental, because it aims to facilitate memory work by establishing a relationship between the viewer, the monument, and the history that is being represented.

Figure 3: Barbed wire and a watchtower at Grūtas Park. Adriao, Wikipedia Commons.

Beyond the history, I assert that the anti-monumental nature of representations at Grūtas Park is also represented by the memory discourses surrounding Soviet history. The subject of Soviet history becomes inherently anti-monumental due to unresolved tensions in the discourses surrounding it, and the consequential difficulty in building affirmative and glorifying monumental works according to the traditional framework of monumentality. It is helpful to compare the memory discourses of the Holocaust and Soviet communism, as both periods are marked with tragic events, conflict, and a difficulty in grappling with the memory of their histories. Rather than drawing similarities between the memory discourses by comparing the totalitarian regimes themselves, it is productive to analyse the ways they are remembered in different geographical and cultural contexts of Eastern and Western Europe. This will help illustrate the tension in regard to Soviet history in Lithuania, clarifying its status as an anti-monumental subject.

In the Western European context, memories of the Holocaust and the Soviet era can be understood as 'hot and cold memory' respectively (Turai 99). Historian Charles S. Maier proposes the 'hot and cold' framework as a distinction between active and passive relations to memory. Hot memory affects the present, demanding continuous engagement with the past (99). It necessitates active memory work, as people are attentive both to the memory itself, as well as the changes in discourse relating to the past throughout the generations (99). Moreover, hot memory may provoke strong emotions and can serve as a basis for political mobilisation (99). In contrast, cold memory remains in the past, somewhat forgotten and not 'worked through' (99). Relation to the past remains passive and does not evoke an emotional response (99).

To some extent, hot and cold memory is similar to Aleida Assmann's distinction of canon and archive. Canon defines active remembering, which means that memory is regularly interacted with and dispersed within society (A. Assmann 395). In this case, the past is remembered as the present. In opposition, the archive is a form of passive remembering, meaning that there is little active engagement with the memory (335). The past is remembered as the past, and so the memory is archived, put away. To explain this distinction, Assmann uses the example of a museum (335). The museum places its most prestigious objects in exhibition halls and hosts exhibitions to present them to viewers. At the same time, the museum also stores objects that are not currently on display and accessible to the viewer (335). The canon is constituted of carefully selected objects presented to the public, while the archive consists of objects that are stored away, but may at any time be added to the museum's exhibition. The active and the passive remembering of the canon and archive may also be understood as a 'working memory' and a 'reference memory' (336). Hot memory functions as canon, always available and influential in societal discourses. In opposition, cold memory is more similar to the archive, as

it is hardly present in societal discourse, and may often be forgotten.

Within this framework, the memory of the Holocaust is a hot memory in the West, for it elicits discussions on remembrance, and produces emotional responses. The memory of Soviet communism is a cold memory, most likely due to the fact that the experience of communism in the West has been peripheral, thus making it difficult for people to empathize with a history that was not experienced directly (Turai 99). Maier states that while portraits of Lenin or Marx are an acceptable piece of furniture in a professor's office, a portrait of Hitler would be completely unthinkable, even in an ironic context (99). This distinction is important for Lithuania and the case of Grūtas Park, as it informs the highly emotional responses to representations of Soviet memory and the controversial reception of Grūtas Park, especially in terms of its ironic portrayals of Soviet history. In contrast to the Western European context, in Eastern Europe, and especially Lithuania, the memory of communism is a hot memory (99). The Soviet occupation elicits strong emotional responses in the East, as the Holocaust does in the West, hence, portraits of Lenin or Marx could be a controversial statement in a professor's office in Eastern Europe. In Lithuania, Soviet memory as a hot memory is best exemplified by the continuous debates concerning Soviet memory, most visible in discussions about the fates of Soviet monuments in public spaces across Lithuania.

In 2015, Soviet era statues were removed from the Green Bridge in Vilnius amidst controversy and debate ("Žaliojo tilto skulptūrų"). The quartet of statues represented social classes within Soviet society, such as students, industrial and agriculture workers, and soldiers (Baločkaitė 2). The statues had aesthetic and artistic value, as they were made by Lithuanian artists and viewed as decorative elements of the Green Bridge ("Žaliojo tilto skulptūrų"). Moreover, they represented everyday citizens of Soviet Lithuania. However, they were commissioned and erected by the Soviet state, acting as symbols of the resented Soviet regime, as well as conveying ideological messages about the ideal structure of society (Baločkaitė 2). Some suggested that the statues should be preserved due to their aesthetic value, regardless of the fact that they were produced in the Soviet era (4). Others argued that the statues were a symbol of totalitarianism, and could possibly have traumatizing effects by acting as constant reminders of a harsh history (4). Officially, the statues were removed for renovation purposes. Yet, discussions surrounding the statues' future – whether they should be reinstated on the bridge, destroyed or exhibited in a museum - did not reach a consensus (5).

Notably, Malinauskas had offered to take the Green Bridge statues and exhibit them at Grūtas Park, however, the municipality refused (Grigaliūnaitė, "Sovietinio režimo simbolis"). The park was the chosen solution for the issue of monuments in the 1990s, but monuments that remained in public spaces and have become a problem recently are not instinctively moved to Grūtas Park. The municipality did not provide an official explanation for their refusal, and the statues remain in the city municipality's storage to this day. As the case of the Green Bridge statues indicates, Soviet history is still a hot memory in Lithuania today. Arguably, it may also indicate that Grūtas Park is not affirmative of a single narrative, hence leading the municipality to refuse the park's offer. Attempting to combine anti-nostalgic and nostalgic sentiments, Grūtas Park could not establish a single narrative for the Green Bridge monuments. Therefore, it would not provide closure to the debates surrounding the statues, which is perhaps not the municipality's intention. The continuous relevance and controversy of Soviet memory in Lithuanian society today, provides a context for representations at Grūtas Park, which determines a multidirectional approach to understanding the memory for the visitors. Thus, the subject is anti-monumental not only due to its painful themes and difficult history, but also its status as a hot memory.

4.2 Visitor Experience: A Carnival of Parody and Irony

Traditional monuments and anti-monuments evoke contrasting emotions in the viewer, appealing to different senses. Traditional monuments typically create an atmosphere of solemnity, towering over visitors and demanding deference (Stevens et al. 727). Primarily, traditional monuments appeal to the sense of sight, while other senses are avoided, or specially forbidden, such as touch (727). Moreover, it is customary to keep a distance from such monuments, appreciating them from afar (727). In contrast, anti-monumental forms challenge traditional monumental relationships, encouraging close, bodily interactions between viewers and monuments (728). Anti-monuments welcome physical interaction, suggesting a direct and horizontal relationship between viewer and monument. As a complex of attractions, Grūtas Park explores diverse sensory experiences. Yet, its interaction with monuments remains merely visual, and somewhat distant. Since the subject of the park's represented memory is complex and multifaceted, Grūtas Park designates different approaches to separate sections of the park. Representing memories of daily life, the park utilizes the sense of taste, sight and hearing to produce nostalgic and parodic experiences. More serious memories, such as deportations and occupation, retain more traditional, visually oriented representations.

At Grūtas Park, visitors may get a taste of the Soviet past. At the café, an assortment of 'Soviet dishes' awaits: a starter of *Nostalgia* in the form of a beet soup, a main dish of meat patties *Goodbye, Youth!* and *Vodka USSR* to wash it all down (Lankauskas 40). The experience concludes with a waiter, adorned with a red neckerchief of the pioneers – a youth organisation in the USSR, bringing the visitor their check (40). For Lankauskas, the cuisine at Grūtas

Park acts as a "catalyst of practical nostalgia ... with a connective purpose" (40). The visitor is connected to the Soviet past through a sensory experience of taste and sight, enveloped in an atmosphere of nostalgia. After lunch, visitors immerse themselves in an audio-visual adventure at a replica of a Soviet Culture House. The Soviets built such buildings all across Lithuania, both in urban and rural locations, to provide spaces for recreational activities (Drėmaitė). At the same time, these spaces encouraged community building according to Soviet societal norms and expectations. The building at Grūtas Park holds a vast collection of documents and photographs, depicting Soviet crimes (Williams 188). It also exhibits an overwhelming amount of various Soviet knick-knacks: medals, pins, etc. From the speakers, visitors may listen to Soviet marching songs or readings of propaganda texts (188). The experiences at the café entail a sense of ironic playfulness, and arguably, they embody the qualities of the carnivalesque.

Mikhail Bakhtin developed the concept of the carnivalesque by analysing narrative frameworks in literature, but its imagery may be found in other mediums and contexts as well. The carnival encompasses laughter, chaos, and most importantly for the case of Grūtas Park, a shifting of order (Lachmann et al. 123). "Carnivalistic laughter is directed toward ... a shift of authorities and truths, a shift of world orders" (123). Hierarchies are overturned, that which was concealed from the eye or repressed is brought into being, and the profane and blasphemous prevail (142). For Bakhtin, the carnivalesque is a temporary respite from the 'established order' and the 'prevailing truth,' marking "the suspension of all hierarchical rank, privileges, norms and prohibitions" (Bakhtin 10). At Grūtas Park, leaders of the Soviet regime and archetypes of Soviet social groups return as characters in live mock-dramatizations (Williams 188). Actors don the roles of Lenin and Stalin, Secretaries of the Central Committee, and members of Soviet youth organisations such as the Soviet Young Pioneers (188). There is something sacrilegious and offensive about bringing back the figures of a regime that desecrated Lithuania in amusing acting performances. Yet, their presence at the park ameliorates the historical pain by mocking and satirizing its criminals. Watching actors give ridiculous ideological speeches both devalues Soviet ideology, deterring visitors from believing in its promises, and simultaneously places visitors in a position of power, able to judge the silly claims of a decrepit regime. The Soviet figures are to be laughed at, mocked, establishing a carnivalesque norm, where the Lithuanian citizen towers above the Soviet system.

The café and the Culture House are also part of the carnivalesque framework. They carry a sense of nostalgia, but at the same time, everything they advertise is shrouded in an ambience of irony. The names of the dishes are somewhat silly and amusing, the sheer amount of propaganda objects produced and now exhibited at the Culture House seems excessive and ridiculous from a modern perspective. However, what allows parody to take effect at the park, and goes beyond Bakhtin's carnivalesque, is the fact that the Soviet period is in the past. The Soviet regime and its ideology failed, and the citizens of present Lithuania look back on it as something not to be repeated. Thus, the carnival at Grūtas Park is not a temporary respite, but rather a cathartic repudiation of the restrictions and oppression of the Soviet era. The anti-monumental is created through the interaction with different senses and the stripping of authority of the Soviet regime through carnivalesque parody.

The park's Soviet monuments, however, do not seem to fit within this framework. Scattered in the forest along a pathway, some distance away, the monuments offer a visual experience, typical of any museum or exhibition. For some monuments, information plaques provide the names of authors and titles of the sculptures, or the locations where they once stood. Yet, the presentation of monuments remains simplistic, the relationship between the viewer and the monument is distant. This choice to keep more traditional forms of presentation is understandable in such sections of the park that are more overtly traumatic, for example, the exhibit of the freight-train car and the barbed wire. Here, identifying the line between ironic subversion of expectations and mocking of trauma, is rather difficult, if not impossible. However, the monuments do not necessarily evoke intense emotional responses, as for instance, they may represent vague ideas about Soviet social groups. Therefore, the monuments are lacking in their representational forms and diverse visitor experiences. This is regrettable, as the carnivalesque could provide interesting counter-monumental interactions between visitors and monuments. This aspect may be further analysed in the site and meaning sections, where I discuss the park's traditional forms of representation.

4.3 Form: Fragmentation and Power Dynamic

Anti-monumental form is the most clearly visually recognizable feature of counter-monuments. Such monuments oppose 'conventional monumental forms,' instead utilizing different, atypical design techniques and materials (Stevens et al. 723). Traditional monuments highlight the presence of solid materials, they are often light in tone in order to stand out from the urban landscape, and emphasize vertical construction (723). They are built on elevated platforms, which helps accentuate their importance and exceptionality within the architectural landscape. Anti-monumental form is vastly different, featuring abstract forms and voids, representing absence and dark tones, and maintaining horizontal lines (723). Anti-monuments are sunken within the landscape, and may be 'shifted off-axis,' fragmented (723). For this reason, anti-monumental forms may be considered "more negative and more complex than traditional ones" (725). The exploration of anti-monumental form at Grūtas Park is perhaps rather subtle, yet it still produces strong effects.

Figure 4: A Lenin bust at Grūtas Park. Yeowatzup, Wikipedia Commons.

Grūtas Park does not directly engage with form, since it simply houses a collection of Soviet statutes, and any new or additional monumental works have not been built at the park. Nevertheless, the aesthetic decisions made by the founders and curators of the exhibition concerning the presentation of monuments create intriguing effects. The monuments are exhibited as they are – restoration work was often not carried out even for the most damaged monuments (figure 4). Many statues are broken in places, often missing the pedestals on which they stood, as a result of the hasty and somewhat careless removals of the 1990s. This presentation of the statues makes a statement in itself. The poor quality of the statues emboldens the idea that the statues do not have any ideological and political weight anymore. Rather, they are only the broken remnants of the past, and as the Soviet regime crumbled, the physical remnants of it did as well. Moreover, most statues at Grūtas Park were not placed on new platforms. Various busts were simply laid on the ground, often even below eye level. The viewer is instinctively placed in a superior position in relation to the monument, and any sense of authority is removed from the monumental work. As Lankauskas argues, "the gazing visitor-voyeur assumes a position of power vis-à-vis a coercive authoritarian system" through their interaction with the monumental symbols of the regime (38).

One of the clearest examples of such subversion of power in monuments is the Lenin statue, removed from Lukiškės Square. Built in 1952, the statue once stood on an imposing pedestal in the centre of the square (fig. 5). The large size of the statue itself in addition to the pedestal, emphasized the verticality of the monumental work. Moreover, the walkways for pedestrians within the square all converged in the centre, where the statue stood, emphasizing its importance and centrality. During the statue's removal in 1991, its legs were broken off. It was eventually moved to Grūtas Park without the pedestal, and was placed directly on the ground (fig. 6). While the legs were re-attached, the lower right foot of the statue remains broken. At Grūtas Park, the statue seems to represent weakness, rather than the imposing authority it once had in Lukiškės Square. What is more, the statue is surrounded by a number of Lenin busts. These artworks are also placed directly on the ground, in a semi-circle formation. The presence of these busts further downplays the importance of the Lenin statue, for it is presented as just one of many monumental works of Lenin. This Lenin monument serves as an example of the way Grūtas Park recontextualizes its exhibits, employing anti-monumental form through the emphasis on horizontality, fragmentation, and subversion of power vis-à-vis the relationship between the viewer and the monuments.

Figure 5: Lenin statue at Lukiškės Square. KTU ASI archyvas.

4.4 Site: Discursive Periphery and Historical Victimhood

Traditional monuments are meant to be seen and visited. Their emphasized verticality towers over citizens and their central and often enclosed position in squares denotes their importance (Stevens et al. 726). Various traditional memorial spaces become tourist destinations, spaces for gatherings and commemorations or military parades. Anti-monuments seek to fulfil other purposes, and do not share similar characteristics (726). Quite often visitor interactions with anti-monuments are accidental, chance encounters during everyday movement throughout a landscape, whether urban or rural (726). Anti-monumental sites invite the viewer to delineate their own approach and relationship to the monument, since they "presume no focused, idealized viewing position" (727). For example, the *Monument Against Fascism* by Esther Shalev-Gerz and Jochen Gerz stood in a commercial neighbourhood in Hamburg, attracting local residents to interact with it and reflect on the history of the Holocaust

Figure 6: Lenin statue at Grūtas Park. Žygimantas Gedvilas, 15min

(Young, "The Counter-Monument" 274). The monument was a square pillar, with a plaque stating, "We invite the citizens of Hamburg and visitors to the town, to add their names here to ours" (274). After a section of names would be filled, the pillar would be lowered into the ground, until it would completely disappear underground. The monument's location was significant in allowing for visitor engagement with it, as it was placed in a pedestrian shopping mall, a busy area, where "package-laden shoppers could like it or hate it, but they could not avoid it" (274). Thus, the site of the monument would compel locals to constantly acknowledge the monument and its memory, while its changing appearance due to the consistent lowering of the pillar, would keep their attention.

In Neukölln, Germany pedestrians passing by the Sonnenallee would be confronted with a visual installation, exposing them to the cruel history of a concentration camp within the space they were currently inhabiting (285). There are no visual remnants of the concentration camp in the area, however, by walking through, pedestrians would trigger a light projection of a text explaining the history of the site's 'now-invisible past' (285). Furthermore, the text would change according to time of day, season and climate, so that the encounter with the memory would continue to be unique (286). Thus, the unique sites of counter-monuments aim to interact with the viewer, while also creating new experiences and subjectivities.

Grūtas Park acts as a traditional monumental space – a tourist destination, located some distance away from main cities in Lithuania. Advertisements of the park highlight its many features and attractions in addition to its monuments. The park's café, playground, and small zoo entice visitors to plan a day visit to explore everything the park has to offer (*Grūto Parkas*). The park is located peripherally, outside of the city, making it less convenient for people to visit. Therefore, the park focuses on advertising various aneminities, rather than just the Soviet monuments and memorabilia exhibition, to invite visitors. However, the park's site also contributes to its failure to integrate into public discourses surrounding Soviet history and memory. Its peripheral location extends to its peripheral status in the public sphere. To some extent, this may inform the reasons why the debates over the fate of Soviet monuments in public discourses continue to cause controversy, despite solutions such as Grūtas Park already being present. If the park aims to serve not only as a representational space, but also encourage introspection, societal debate and communication between citizens, it has not truly succeeded in doing so.

Within the counter-monumental framework, interaction with monuments and their representational history is an often accidental, daily occurrence, since the goals of counter-monumentality include the facilitation of continuous memory work. This is difficult to imagine at Grūtas Park, as the park is simply not within reach in the everyday lives of people. Williams explains that the location of the park might have been chosen in order to "visibly devalue their [Soviet monuments] former political capital" (190). This is indeed what the positions of the Lithuanian government and founder Malinauskas would suggest, as the primary problem concerning the monuments were their ideological and political meanings, which became unacceptable in independent Lithuania. However, physically distancing the monuments also results in a mental distancing – the Soviet period is perceived as a thing of the past, not prompting any further considerations or analysis (193). Williams argues that the park may actually be "divesting local visitors of the task of creating greater civic dialogue around questions of historical guilt, culpability and passivity" (195). This would prove detrimental within a counter-monumental framework, as it aims to stimulate personal interactions with monuments, emphasizing their referential and changing meanings. Nevertheless, it may be worthwhile to determine what other meanings the park's occupied space may carry.

Building Grūtas Park in Grūtas forest, where Lithuanian partisans laid down their lives was controversial and considered a mockery of the site's history. For many, encountering Soviet ideological symbols in such a space was simply unacceptable (Williams 193). However, I would suggest that the park endorses the site's history, affirming Lithuanian national historical narratives in a manner typical of traditional monuments. The park includes the spatial context within its primary representational meaning – a critique of the Soviet regime. Visitors are provided with contextual and visual clues, guiding them towards a singular interpretation. The location has a painful history, which is both a source of sadness and pride in the nation, as it alludes to the unsuccessful fight for independence. The park's website informs visitors that the park is a critique of the Soviet era and that it exposes the deceit of Soviet ideology (*Grūtas Park*). It affirms the ideas of partisan history, fighting Soviet ideology on intellectual terms. First entering the park, the visitor notices the exhibits of the history of deportations, which attains thematic links with the fight for independence by generating feelings of victimhood and loss. Therefore, the park's site explores traditional monumental forms, as it conveys a particular narrative of struggle and victimization, and affirms Lithuanian identity.

4.5 Meaning: A Nationalistic Narrative and Ambiguity

Monuments are usually built with the purpose of representing and conveying specific meanings. Traditional monuments are affirmative and didactic, representing specific ideas and messages (Stevens et al. 728). They utilize 'figural representations,' make 'textual or graphic references' to specific, somehow historically or culturally significant events, people or locations (728). They may also engage with "allegorical figures and archetypal symbolic forms" (728). Anti-

monumental forms of representing meaning, on the other hand, avoid making claims or imparting specific knowledge (728). They are ambiguous, carrying various possible interpretations that depend on the visitor's background or information made available by institutions where the monuments may appear, such as museums (728). Using abstract forms, for instance, anti-monuments can avoid imparting set meanings in order to transfer the burden of memory work onto the viewer (728). As indicated by the specific narrative representation of the site, Grūtas Park leans towards imparting a particular meaning of critique through its exhibits. However, different stakeholders have insisted on the meaning the park represents, affecting the ways in which the park itself understands and presents its meaning.

The project for Grūtas Park was first proposed in order to ensure the preservation of Soviet history and memory, as well as provide a new space and purpose for Soviet monuments. The official intended meaning of the park's exhibition stems from these motivations. Founder Malinauskas asserts that the park welcomes visitors to teach them about Soviet history ("2001"). For him, the framing of the exhibits at the park is both critical and deterring: "Our task is to constantly caution future generations about the consequences of the consolidation of dictatorships" ("2001"). The park represents the responsibility of historical accountability and the duty to not repeat past mistakes. Official information sources also state that the exhibition exposes the failures and immorality of Soviet ideology, as well as critiques its negative impact on society (*Grūto Parkas*). The park delivers on such promises, as it documents the history of Soviet crimes, such as deportations, critiques and mocks the Soviet social order in informational texts and theatre performances. However, while the park claims to represent a unified message, it simultaneously encapsulates various ideas and perceptions. The controversy surrounding the park indicates that even though a number of people believed in the intended message, they disputed the park's dedication to representing it. They noticed messages of positive Soviet nostalgia and problematized them.

The carnivalesque and the ironic, the light-hearted experiences at the café or the Luna Park, carve out a space in the park's memory discourse for other interpretations of the Soviet era. They draw attention to youth and daily life, constructing an experience of nostalgia and innocent naivety. These sections of the park do not necessarily engage with an ideological critique of the Soviet period. They are a part of the Soviet experience in Lithuania, yet they do not evoke contemplation about the political and ideological aspects of the Soviet system. A dissonance between the official presentations of the park's meaning and the experienced reality at the park escalates the discursive controversy surrounding Grūtas Park. Were the park open about its ambiguous meanings and uncertainties, and embraced a multi-faceted meaning to its exhibition, it may not have garnered such intense criticism. The park engages with anti-monumental notions of ambiguous meanings and the importance of individual interpretations and multidirectional memories to a larger extent than the park's advertisements and organisational body maintains. Denying other possible meanings, whether to avoid social criticism or concede to governmental intentions, limits the freedom of discursive practices at Grūtas Park. The visitor is not given the opportunity to simply experience the exhibition and draw their own conclusions. The answer is provided for them, even if it does not embody the entirety of the park's experience, nor the entirety of the Soviet experience in Lithuania.

Grūtas Park attempts to represent a multidirectional memory, a memory that stretches into different directions, and references multiple contexts. However, it does not embrace this memory nor does it provide a space for interpretation. The reasons for such limiting may be many: political, legal, or social. On a governmental level, the state deems it necessary to establish a Soviet history that denies revisionism and uncritical glorification of the period, which is common to Soviet sympathisers, as well as Lithuania's political opponents such as Russia. The state places many restrictions on narrations of Soviet history, such as a ban on all Soviet symbols from appearing in public spaces (Bungardaitė & Radauskas). They are permitted in educational contexts, which explains their presence at Grūtas Park. However, transferring them into public spaces is still a legal and even moral concern, as was shown in 2018. Wishing to advertise Grūtas Park at an international tourism exhibition, employees brought leaflets and various trinkets to their stand, which featured Soviet symbols (Bungardaitė & Radauskas). The act caused a public outrage and an investigation into the legality of the employees' actions (Bungardaitė & Radauskas). The stand was swiftly rid of any Soviet symbols, and the municipality of Druskininkai issued an apology to the exhibition's organisers. Hence, there are limits to the freedom of expression in public spaces that make persons wary of pushing the boundaries. Albeit, Grūtas Park already enjoys a degree of freedom in expression, which suggests that it could go beyond its current, still somewhat restricted discourse, yet chooses not to.

Finally, it is worthwhile to consider whether Lithuanian society itself does not create obstacles for moving towards more liberal and wide-ranging interpretations of Soviet history. After all, Grūtas Park garnered a large amount of criticism from vocal nationalist and patriot groups, as well as government representatives. But I have also indicated that a part of society wishes to have the freedom to relate to nostalgic, positive memories of the Soviet era, without being marginalised (Klumbytė, "Memory" 295). Even if they may not be the majority, they establish a precedent for a reconsideration of the monolithic representations of the period. Shaping such a transition to an open discourse may not be simple, however, this does not suggest that action should not be taken. Nevertheless, currently, Grūtas Park is to an extent embracing the anti-monumental notion of meaning, as it encapsulates different experiences and interpretations of the

Soviet period, from hatred and pain to nostalgia and laughter. However, whether due to governmental or societal expectations, the park itself constricts its represented meaning to a singular interpretation, consequently limiting the potential of discursive interactions and personal visitor contemplation at the park.

5 Facilitating a Dialogue with the Soviet Past

Grūtas Park appropriates anti-monumental forms by representing the hot memory of Soviet communism, creating a carnivalesque visitor experience, and subverting visitor and monument relationships by curating the park's exhibits. At the same time, it incorporates traditional monumental forms by attempting to present an undivided memory narrative and disengaging from public discourse due to its peripheral location. However, it is becoming clear that the multidirectional memory of the Soviet era necessitates a diversified approach to representation, and motivates the usage of the counter-monumental framework. With aspects, such as the presence of nostalgic discourses in society, Grūtas Park is already struggling to integrate differing narratives without dismantling its own narrative of a unified history of Soviet memory. Further investigating the park through the second approach to counter-monumentality – the dialogic framework, it becomes evident that more extensive commitment to a counter-monumental form of representation is a productive asset to the park's memory discourse. It can cohesively contain the different strands of meaning at Grūtas Park and it may also help integrate the park into public discourses surrounding Soviet memory.

The dialogic monument provides a comparison and a juxtaposition with an existing monument (Stevens et al. 729). Its purpose is to question or critique the values and meanings of the existing monument (729). As Stevens et al. argue, "A dialogic coupling dramatizes new meanings beyond those conveyed by each of the works considered individually" (729). A dialogic monument, therefore, must be spatially linked to another monument to allow for a comparison, and it must contain historical and representational specificities (729). A dialogic approach seems promising in the case of Lithuania, since the government expected direct communication with already existing Soviet monuments. They deliberated over the appropriate form of commemoration by recontextualizing and repurposing ideologically improper monuments. Yet, despite the park's apparent suitability for developing monumental dialogue, a discursive conversation did not begin at Grūtas Park.

Grūtas Park emulates a museum space for Soviet monuments. Here, they become exhibits, curated and appraised according to the founder's ideas. New monumental forms are not built at the park, and a dialogue develops solely between the visitor and the monuments. The need for an inter-monumental dialogue in the first place can be indicated by the continuing tensions surrounding the city squares, streets, and alleys, where Soviet monuments once stood. The statues themselves have been removed, yet their historical and memorial impact continues to haunt these spaces. The Lithuanian state is attempting to fill in the memory voids the Soviet monuments left, replacing them with new memories and narratives. However, this process is problematic and contentious. In Lukiškės Square, for instance, a fitting replacement for the Lenin statue has not been found to this day. As a temporary solution, a flagpole with the Coat of arms of Lithuania (*Vytis*) stood at the centre of the square. Feeling an obligation to design a memorial space, the government of Lithuania held a contest for design projects in 2017. Andrius Labašauskas won with the monument project *Laisvės Kalva* (*Freedom Hill*), which memorializes freedom fighters of the partisan war (Grigaliūnaitė, "Sovietinio režimo simbolis"). Yet, the decision was controversial, and a model of the new monument was vandalized with red paint (Grigaliūnaitė, "Sovietinio režimo simbolis"). Not everyone believed in the idea of *Laisvės Kalva*, with the most popular alternative solution being the preservation of the *Vytis* symbol in some form (Grigaliūnaitė, "Sovietinio režimo simbolis"). To some extent, the conflict over Lukiškės Squares stems from the fact that many people imagine different ideas for commemoration, highlighting different parts of Soviet history as most deserving of memorialization.

The implementation of the project remains on hold, in large part due to its mixed reception by the public. An effort is being made to counteract the heavy weight of the ideological and political past the monuments have left in public spaces. However, a much-needed interaction with the past is not taking place. Rather than acknowledging the history of the square and reimagining its meanings through new monumental works, monuments are used to negate the past by obscuring it with representations of other historical narratives. Simply removing a monument and replacing it with a new one is not sufficient to shift the memory meaning of the square, as it is a *lieu de mémoire*, as defined by Pierre Nora (Nora 7). The square acts as a material and functional space, which carries symbolic meaning that references the historical past (19). Simply put, many people still remember when the square was called Lenin Square and boasted a monument of Lenin, which motivates a need to counteract these memories with a new monumental work (Grigaliūnaitė, "Sovietinio režimo simbolis"). However, progress cannot happen, if the past of the square is not addressed and worked through. A solution to the disagreements at Lukiškės Square may also be a solution for the deficiency of memory work at Grūtas Park. It is necessary to establish a dialogue with the past, a dialogue between Grūtas Park and its monuments, and the many spaces where these monuments once stood.

The Lenin statue at Grūtas Park features an informational plaque, which describes where the statue once stood

and when it was removed. It speaks directly to the monument's previous context and its impact. A similar dialogue could commence at Lukiškės Square as well, explaining why the monument was removed and why it resides at Grūtas Park today. Similar interactions between monuments and public spaces in cities could be applied to other statues at Grūtas Park. Doing so, would integrate Grūtas Park into the public discourse, despite its peripheral site, constructing a memory network. Building such connections could help explore the contextual and referential nature of Soviet memory as a multidirectional memory, facilitating more intimate and fulfilling experiences both at Grūtas Park and in urban spaces in Lithuania. Both the anti-monumental dimensions and the dialogic possibility reveal the vast potential of Grūtas Park as a counter-monumental space. Further exploring possibilities of representation at the park in the future, could reveal new subjectivities, intricacies and references of the multidirectional memory of Soviet Lithuania.

6 Conclusion

Grūtas Park represents the complicated and multidirectional memory of Soviet Lithuania. Narratives of anti-nostalgia and nostalgia, which interact and contend with one another in public discourses, are referenced in the various attractions and exhibits at the park. Seeking to represent both narratives, Grūtas Park employs various forms of representation, both counter-monumental and traditional. This paper reveals that Grūtas Park successfully engages with the counter-monumental framework and its forms by representing a difficult history that is still a hot memory in Lithuania today. It creates fascinating and amusing visitor experiences by parodying and mocking figures of the Soviet regime and the idealized norms of Soviet society. Constructing a carnivalesque image of Soviet memory, the park subverts hierarchies and diminishes the political and ideological influence of the Soviet regime. It ameliorates the pain and heaviness of the history by placing the visitor in a position of power and strength against the Soviet state in its symbolic forms. Embracing the fractured and crumbled Soviet monuments, Grūtas Park further lessens their ideological and political weight by utilizing counter-monumental forms, and for instance, placing emphasis on horizontality.

Simultaneously, the park strives to realize its initial statement of purpose from the 1990s. It aims to act as a reminder of Soviet history, caution future generations against repeating past mistakes, and store the no longer needed symbols of a former regime. Therefore, its peripheral location devalues Soviet ideology, but it also distances the park from public spaces and discourses. By claiming to represent a single unified narrative, the park fails to recognize its own ambiguous meanings, encountered through its diverse representations of anti-nostalgic and nostalgic sentiments, as well as different aspects of Soviet history such as deportations and daily life. Observing the lack of dialogue between the Soviet monuments and their previously occupied spaces, which obtain aspects of *lieux de mémoire*, this paper suggests that Grūtas Park has an unrealized potential for truly embodying the notions of counter-monumentality. Facilitating a dialogue between the park's exhibits and public spaces in Lithuania may help address the current memory debates happening in Lithuania, as well as ensure the continuous memory work concerning Soviet history, by reintegrating Grūtas Park into public discourse. Overall, the park offers a unique and rich experience of counter-monumental representations of a multidirectional memory. If the founders and curators of the park further develop the connections between the monuments and public spaces, the park may reach new potential in representing history, allowing visitors to have truly introspective and thoughtful experiences at Grūtas Park.

Works cited

Antanavičius, Ugnius. "Kaip sovietinė valdžia perkūrė Lietuvos himną ir kodėl „Tautiškos giesmės" atsisakyta ne iškart." *15min*, 15 Jul. 2020, www.15min. Accessed 23 Oct. 2021.

Assmann, Aleida. "From Canon and Archive." *The Collective Memory Reader*, edited by Jeffrey K. Olick et al., Oxford University Press, 2011, pp. 334-337.

Assmann, Jan. "Collective Memory and Cultural Identity." *The Collective Memory Reader*, edited by Jeffrey K. Olick et al., Oxford University Press, 2011, pp. 212-215.

Bakhtin, Mikhail M. "Introduction." *Rabelais and His World*, translated by Helene Iswolsky, Indiana University Press, 1984, pp. 1-58.

Baločkaitė, Rasa. "The New Culture Wars in Lithuania: Trouble with Soviet Heritage." *Cultures of History Forum*, 2015, pp. 1-14. doi:10.25626/0034. Accessed 13 Feb. 2021.

Bungardaitė, Indrė and Gintaras Radauskas. "Tarptautinę turizmo parodą Vilniuje apkartino komunistiniai simboliai iš Druskininkų." *15min*, 27 Jan. 2018, www.15min. Accessed 1 Jun. 2021.

Černiauskaitė, Lina and Elena Jasiūnaitė. "Sovietinės šventės – ideologijos spektaklis. Pokalbis su istoriku Arūnu Streikumi." *15min*, 17 Dec. 2017, www.15min. Accessed 25 Apr. 2021.

Dapkus, Liudas. "Stalin's world." *The Spokesman-Review*, 14 May 2006, `https://www.spokesman.com/stories/2006/may/14/stalins-world/`. Accessed 25 Mar. 2021.

Davoliūtė, Violeta. The Making and Breaking of Soviet Lithuania: *Memory and Modernity in the Wake of War*.

Routledge, 2013.

Drėmaitė, Marija. "Vėlyvojo Sovietmečio Grimasos." *MO Muziejus*, www.mmcentras. Accessed 2 Jun. 2021.

Erll, Astrid. "The Invention of Cultural Memory: A Short History of Memory Studies." *Memory in Culture*, translated by Sara B. Young, Palgrave Macmillan, 2011, pp. 13-37.

Grigaliūnaitė, Violeta. "Gegužės 9-ąją Vilniaus Antakalnio kapinėse – raudonų gvazdikų ir Georgijaus juostelių jūra." *15min*, 9 May 2018, www.15min. Accessed 11 Apr. 2021.

Grigaliūnaitė, Violeta. "Sovietinio režimo simbolis – Lenino paminklas – Vilniaus Lukiškių aikštę paliko be kojų ir piršto." *15min*, 11 Mar. 2020, www.15min. Accessed 5 Mar. 2021.

Grigoravičienė, Erika. "Paminklai Idėjoms, Žmonėms, Formoms: XX. a. 7-8 Dešimtmetis." *MO Muziejus*, http://www.mmcentras.lt/kulturos-istorija/kulturos-istorija/daile/skulptura/paminklai-idejoms-zmonems-formoms-xx-a-78-desimtmetis/79476. Accessed 5 May 2021.

Grūto Parkas. http://grutoparkas.lt/. Accessed 8 Feb. 2021

Gussi, Alexandru. "On the Relationship between Politics of Memory and the State's Rapport with the Communist Past. " *Remembrance, History, and Justice*, edited by Vladimir Tismaneanu and Bogdan C. Iacob, Central European University Press, 2015, pp. 125-142

Isaac, Rami K., and Laurencija Budryte-Ausiejiene. "Interpreting the Emotions of Visitors: A Study of Visitor Comment Books at the Grūtas Park Museum, Lithuania." *Scandinavian Journal of Hospitality and Tourism*, vol. 15, no. 4, Routledge, 2015, pp. 400-424, doi:10.1080/15022250.2015.1024818. Accessed 14 Apr. 2021.

Klumbytė, Neringa. "Dabarties istorijos paraščių žmonės: atsiminimai apie sovietmetį ir kasdienybės patirtys kaimo bendruomenėse." *Politologija*, vol. 3, no. 35, 2004, pp. 1-21.

Klumbytė, Neringa. "Memory, Identity, and Citizenship in Lithuania." *Journal of Baltic Studies*, vol. 41, no. 3, 2010, pp. 295-313. JSTOR, www.jstor.org/stable/43212969. Accessed 8 Feb. 2021.

Kuczyńska-Zonik, Alexandra. "Dissonant Heritage: Soviet Monuments in Central and Eastern Europe." *Historical Memory of Central and East European Communism*, edited by Agnieszka Mrozik and Stanislav Holubec, Routledge, 2018, pp. 141-168.

Lachmann, Renate. "Bakhtin and Carnival: Culture as Counter-Culture." Translated by Raoul Eshelman and Marc Davis. *Cultural Critique*, no. 11, 1988-1989, pp. 115-152. JSTOR, https://www.jstor.org/stable/1354246. Accessed 23 May 2021.

Lankauskas, Gediminas. "Sensuous (Re) Collections: The Sight and Taste of Socialism at Grūtas Statue Park, Lithuania." *The Senses and Society*, vol. 1, no. 1, 2015, pp. 27-52. doi:10.2752/174589206778055682. Accessed 11 Feb. 2021.

"Lietuvos katalikų bažnyčios kronika." *Lietuvos Katalikų Bažnyčios Kronika*, http://www.lkbkronika.lt/index.php/apie-kronikas2/3698-lietuvos-katalik-baznycios-kronika. Accessed 4 May 2021.

Makhotina, Ekaterina. "Between heritage and (identity) politics: dealing with the signs of communism in post-Soviet Lithuania." *National Identities*, 2020, doi:10.1080/14608944.2020.1784123. Accessed 8 Feb. 2021.

Marciniak, Katarzyna. "Lenin in Los Angeles: Counter-Memories, Recycling Socialism." *The Oxford Handbook of Communist Visual Cultures*, edited by Aga Skrodzka, Xiaoning Lu, and Katarzyna Marciniak, doi:10.1093/oxfordhb/9780190885533.013.32. Accessed 10 Feb. 2021.

Matonytė, Irmina. "The Elites' Games in the Field of Memory: Insights from Lithuania." *History, Memory and Politics in Central and Eastern Europe*, edited by Georges Mink and Laure Neumayer, Palgrave Macmillan, 2013, pp. 105-120.

"Melagių dieną oficialiai atidarytas Grūto parkas." *Delfi*, 2 Apr. 2001, https://www.delfi.lt/archive/melagiu-diena-oficialiai-atidarytas-gruto-parkas.d?id=258193. Accessed 25 Mar. 2021.

Michnik, Adam, and Elzbieta Matynia. "Nationalism." *Social Research,* vol. 58, no. 4, 1991, pp. 757-763. JSTOR, www.jstor.org/stable/40970671. Accessed 24 Mar. 2021.

Misija Sibiras. https://misijasibiras.lt/. Accessed 24 Mar. 2021.

Nora, Pierre. "Between Memory and History: Les Lieux de Mémoire." *Representations*, no. 26, 1989, pp. 7-24. JSTOR, https://www.jstor.org/stable/2928520. Accessed 2 Jun. 2021.

"Parko Gimimas ir Mirtis." *Procesai*, written by Jolanta Marcišauskytė-Jurašienė and Mikas Žukauskas, LRT, 2021.

"Prie ruso buvo geriau ir konservuoti ananasai || Laikykitės ten su Andriumi Tapinu || S03E25." *YouTube*, uploaded by LaisvėsTV, 17 March 2019, https://www.youtube.com/watch?v=qdGl9bqgEIs.

Renan, Ernest. "What is a nation?" *Nations and Narration*, translated by Martin Thom, edited by Homi K. Bhabha, Routledge, 1990, pp. 8-22.

Rothberg, Michael. "Introduction: Theorizing Multidirectional Memory in a Transnational Age." *Multidirectional Memory: Remembering the Holocaust in the Age of Decolonization*, edited by Mieke Bal and Hent de Vries, Stanford University Press, 2009, pp. 1-29.

Ryžtas žydėti. https://www.birzelio14.lt/. Accessed 24 Mar. 2021.

"Stalin World." SBS, Sydney, 26 Sep. 2010.

Stevens, Quentin, et al. "Counter-Monuments: The Anti-Monumental and the Dialogic." *The Journal of Architecture*, vol. 23, no. 5, Routledge, 2018, pp. 718-739, doi:10.1080/13602365.2018.1495914. Accessed 1 Mar. 2021.

Turai, Hedvig. "Past Unmastered: Hot and Cold Memory in Hungary." *Third Text*, vol. 23, no.1, 2009, pp. 97-106. EBSCOhost, doi:10.1080/09528820902786735. Accessed 19 May 2021.

Williams, Paul. "The Afterlife of Communist Statuary: Hungary's Szoborpark and Lithuania's Grutas Park." *Forum for Modern Language Studies*, vol. 44, no. 2, 2008, pp. 185-198. doi:10.1093/fmls/cqn003. Accessed 11 Feb. 2021.

Young, James E. "At Memory's Edge: After-Images of the Holocaust in Contemporary Art and Architecture." *The Collective Memory Reader*, edited by Jeffrey K. Olick et al., Oxford University Press, 2011, pp. 371-374.

Young, James E. "The Counter-Monument: Memory against Itself in Germany Today." *Critical Inquiry*, vol. 18, no. 2, 1992, pp. 267-296. JSTOR, http://www.jstor.org/stable/1343784. Accessed 19 May 2021.

"Žaliojo tilto skulptūrų galima buvo ir nešalinti: geriau būtų virtusios mikimauzais?" *Veranda*, written by Violeta Baublienė, LRT, 2020.

"2001." *Mūsų laisvės metai*, LRT, 2014.

Figures

Figures 1-2, 6 Gedvilas, Žygimantas. "Sovietinių skulptūrų Grūto Parkas." *15min*, www.15min. Used with permission.

Figure 3. Adriao. "Entrance to Grūto Parkas, Lithuania." *Wikipedia Commons*, https://commons.wikimedia.org/wiki/File:Gr%C5%ABto_parkas_-_Entrance.JPG.JPG.

Figure 4. Yeowatzup. "Grūto Parkas, Lithuania." *Wikimedia Commons*, https://commons.wikimedia.org/wiki/File:Gruto_Parkas,_Lithuania_(7182810437).jpg.jpg.

Figure 5. "Lukiškių aikštė 7 deš. pab." *KTU ASI archyvas*, http://www.autc.lt/lt/architekturos-objektai/745. Used with permission.