Examining Dehumanization Through the "Political Brain Perspective": Towards a Minimal Neuropolitical Theory for Hyperdiverse Societies

Liya Yu

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

Abstract

Examining Dehumanization Through the "Political Brain Perspective": Towards a Minimal Neuropolitical Theory for Hyperdiverse Societies

Liya Yu

What cognitive conditions need to be in place in order for cooperation, and potentially, solidarity, to exist in hyperdiverse societies? What aspects of our social human brain are indispensible when it comes to achieving shared goals in divided liberal democracies? This dissertation singles out the dehumanized perception and categorization of out-groups as one of the most decisive disruptors of political cooperation. I develop an interdisciplinary model – the "Political Brain Perspective" (PBP) – that combines political theory, political science and social neuroscience insights to advance my argument about dehumanization in both domestic and International political contexts.

I argue that dehumanized perception at the brain level is politically troublesome because it disables an important social brain function called mentalizing, which is foundational for both basic political transactions and more complex feelings such as empathy. I show how this is relevant in regard to the neuropolitical duties public representatives owe to their constituents in a diverse liberal democracy, and further, how various liberal traditions such as social contract, multiculturalism and human rights theories have hitherto ignored dehumanization as a fundamental disruptor to any political cooperative process.

At the international level, I examine the potential for dehumanization within civilizational discourses in history, with a particular focus on the post-Cold War distinction between "civilized" and "barbarians". I show that in the international context of genocide, intergroup conflict and identity politics, dehumanizing categories not only diminish the cognitive reasoning and mentalizing abilities of the dehumanizer, but also have an intense impact on the dehumanized, in the form of reciprocal dehumanization and retributive violence.

Based on the epistemological premises of the PBP, I contend that a minimal neuropolitical theory of cooperation ought not to prioritize an ontological concept of human

dignity but instead treat the ascription of humanness as an interpersonal brain mechanism. The brain data, in other words, can only tell us what our brains do when engaging in politics, not who we are as political beings in an essentialist way.

In sum, this dissertation highlights the need for political scientists to pay attention to the neuronal mechanisms underlying dehumanization, and to distinguish it from other forms of exclusion and prejudice as a fundamental brain ability in its own right.

Table of Contents

Acknowledgements

An interdisciplinary dissertation project is always an adventure and a gamble. In order for the gamble to pay off and the adventure to take place without too many bruises along the way, one needs extraordinary intellectual and personal support.

I would first and foremost like to thank my three main advisers, David Johnston and Jack Snyder at Columbia University, and Lasana Harris at University College London. David Johnston's support – from my admission to Columbia's Ph.D. program, to pursuing a neuropolitical dissertation project, to the belief that I had it in me to juggle research and motherhood, and to reading the last pages of my dissertation draft – has been constant and unwavering, through all the highs and lows of this dissertation project. David patiently guided me towards "riding the bicycle" of research framing and kept encouraging me not to give up on our shared belief that Hobbes' insights are highly relevant for today's fractured societies, and need to be transformed into a truly innovative political theory for the 21st century. David's fearlessness and genuine liberal-minded disregard for disciplinary boundaries liberated me in turn to explore the questions that had brought me to pursue graduate studies in the U.S. David instinctually understood where I was coming from and gave me the necessary freedom to explore questions around identity that troubled me most.

Jack Snyder has been my intellectual guardian angel during my time at Columbia, seeing something in me that I did not know existed when I first entered his graduate class on nationalism. Just like David, Jack's intellect is truly free flowing and wide-ranging, but not without demanding the necessary analytic discipline from his students. Jack was always the first one to comment on the countless drafts I circulated over the years, guiding me with the rare and precious mixture of patience, criticism and openness that any student can only dream of in a mentor. In particular, Jack's insights into brain science from a political perspective – continuing to send me relevant articles and dragging me along to interdisciplinary seminars over the years – and his belief in the need for thoughtfully introducing psychological and cognitive perspectives into the analysis of world politics were central to motivating me with my project. Jack always gave me the sense that whatever I do is worthwhile, no matter how I decided to tackle this unwieldy

phenomenon of "identity politics" (which never stopped him from also being one of my most constructive critics). He always made me feel that he is my biggest fan – I am certainly his. It was also refreshing to be able to share other non-academic passions over the years, such as our love for opera and sneezing pandas.

Lasana Harris has been an incredible mentor and source of support during this project. Without his endorsement in his capacity as a social neuroscientist studying exclusion and dehumanization, and his philosophical and humanistic prowess and knowledge, I would not have had the confidence to envision this project in its radicality. Lasana's own unconventional, interdisciplinary way of thinking about social and political phenomena is inspiring in its depth and moral concern, and served as crucial inspiration during many stages of this project when I wanted to give up. I am deeply indebted to him for trusting me from the start – even though he had only met me through e-mail contact – and for serving as a role model for young minority academics of color like myself. In this context I would also like to thank Susan Fiske for her immediate support for my initial project idea and for connecting me with Lasana. Without people like Lasana in academia, I might have let my identity questions bury rather than empower me.

In addition, I would like to express my gratitude to Robert Jervis and Lena Verdeli, who agreed to serve on my dissertation committee. I thank Lena especially for agreeing to serve on my committee on such short notice; her comments at the defense will affect my thinking about this project far beyond the dissertation defense. Andrew Nathan and Ira Katznelson kindly served on my dissertation proposal committee; Andrew patiently commented on various drafts of the proposal.

Andrew Nathan grew to be one of my closest intellectual interlocutors and friends during my time at Columbia. He was unable to lend his wholehearted support to the brain perspective, but he always supported me without compromise. Our heated but fair arguments about political science epistemology and identity politics matured me intellectually, and made me understand that your closest mentors do not always have to agree with you. I believe that what Andrew has done for generations of Chinese scholars and students passing through Columbia – in terms of humanizing them and allowing them to humanize themselves, by enabling them to develop their distinct voice free from civilizationary shame towards the West – is an invaluable gift and service to the Chinese

people, and will be felt by us for many decades to come. Serving as Andrew's teaching assistant and being treated as an equal in advancing the growth of our students helped me find the pedagogical mission of my research project. Relatedly, I am grateful to the Weatherhead East Asian Institute for granting me a fellowship in 2010, and the Chung Wei-Chiu fellowship program.

The intellectual roots of this dissertation can be traced back to my time at the University of Cambridge as an undergraduate. Founding the interdisciplinary "Thinking Society" with Quentin Skinner strengthened me in my belief that cross-disciplinary thinking is worthwhile, and necessary for tackling the big questions of our time. Being trained in the *Cambridge School of Political Thought* gave me the necessary discipline and humbleness in facing up to my re-conceptualization efforts of major political theory ideas in this dissertation. Nicholas James, Duncan Kelly, Harald Wydra, the late Gerard Duveen (who sparked my initial interest in social psychology), Margaret Rigaud-Drayton, Andrew Lovett and John H. Coates were great sources of intellectual and personal support. Melissa Lane, another outstanding mentor, deserves special mention for taking me under her wing, and for continuing to do so during my graduate studies in the U.S. Kelvin Bowkett decided to take a risk with me and admit me to Christ's College based on an interview conducted in Shanghai. Nicola Smith helped me to reclaim my own vulnerable brain, and showed me the importance of brain insights in understanding personal and political identity crises.

My time spent as a research fellow at the *Research Institute for Philosophy in Hannover* (FIPH), Germany was crucial in crystallizing my focus on dehumanization as the main subject of my study. Jürgen Manemann, one of the most impressive intellectuals I have met, taught me that what counted was not whether I would ever be acknowledged by my fellow Germans as one of them, but that humanizing myself was the first priority. Our shared post-War German concern for Jewish identity after the Shoa, as well as for humanizing those who are most vulnerable in society greatly influenced the moral concerns of this dissertation. Similarly, Rolf Elberfeld at Hildesheim University encouraged me to focus on dehumanization and set my intellectual compass in the right direction. I greatly admire Rolf for having transcended "Western" and "Asian" boundaries in his own research, and for instilling in those around him the drive to make

sense of cultural questions in uncompromising humanizing ways. I also am also grateful to Yoko Arisaka and our conversations about race, and her own inspiring work on the philosophy of racism. I thank participants at research seminars at the FIPH, Hildesheim University and the Leibniz University of Hannover (especially Rainer Schmalz-Bruns) for insightful comments and lively discussion.

I greatly benefited from having presented my work at various conferences and venues: I wish to thank participants at the *Asia-Pacific Cultural and Economics Summit* at Hongkong City University (2013), the *Technische University Berlin's "Subjekt und Identität" Conference* (2013), the *New York State Political Science Conference* (2014), Columbia's *Political Theory Graduate Works-in-Progress Seminar* (2015), the *Midwest Political Science Association Annual Meeting* (2015), the *American Political Science Association Annual Meeting* (2015), UVA's *Political Philosophy, Policy and Law Research Seminar* (2016), and my lecture given to *Columbia's Neuroscience Society*. I am particularly indebted to Elaine Denny, Judith Failer, Lutz Fricke, Celeste Lo and Alessandro del Ponte for meaningful conversations and their encouragement of my project. I would also like to thank the formidable Robert Bernasconi for reading through an earlier draft of Chapter 3, his enthusiasm for my project, and for setting the bar of excellence extremely high through his own outstanding scholarly work on race and racism in Western political theory.

At Columbia, other mentors, fellow colleagues and former students generously lent their cognitive resources and time, and gave crucial feedback along the way. I wish to mention in particular Melissa Schwartzberg, William McAllister, Bjorn Gomez, Brett Mayer, Lucas Leeman, Aurelia Bardon, Guido Parietti, Andreas Avgousti, Alex Rigas, Alexander de la Paz, Yuan Yuan (now at Yale), Anna Snyder, Sophie Wilkowske, Max Ridge, and Noelle Martinez. Ashraf Ahmed, the kindest of all colleagues, deserves credit for pointing out that the PBP might redefine "where politics begins".

Solongo Wandan, Yao Lin, Florence Laroque, Irina Soboleva and Justine Guichard were especially close intellectual companions and dear friends during times of triumph and crisis. Justine Guichard in particular was there with her love and wisdom every step along the way. Without their friendship I could not have persisted until the end.

Working with the awe-inspiring Lena Verdeli and her fabulous team of clinical psychologists (especially Bryan Cheng) at *Columbia's Global Mental Health Lab* on our mental health project for Chinese youth has opened my eyes for how my dissertation research can be applied at the intersection of mental health advocacy and education.

At UVa, I have benefited from the Department of Politics' incredible welcome and kind interest in my work. I thank David Leblang and Paul Freedman for letting me teach my course on the political brain, which was modelled almost completely on my dissertation. I am especially grateful to my students in both the Fall 2016 and Spring 2017 semester for stimulating my thoughts during the writing process, and their heartening enthusiasm for the unusual field of "neuropolitics". Colin Bird backed me from the start in every way he could and engaged me in thought-provoking conversations. Paul Morrow generously shared his time, contacts and friendship at the most crucial moment, and provided astute comments to earlier drafts of the dissertation. Sophie Trawalter, Denise Walsh, Lawrie Balfour, Steven White, Jennifer Rubinstein, George Klosko, Cora Schenberg, Jeffery Grossman, Asher Biemann, Rachel Wahl, Caitlin Wylie, Maurice Apprey, Daniel Ciudin, Brandon Ng, Katie Lancaster, Nomi Dave, Ahmed Teleb and Megan Welton expressed interest in my project and tried to include me into the wider UVa community. A special thanks goes to Mark Edwards for promoting my work.

Finally, a dissertation project spanning such a long time is not possible without a strong family and friends support network. I wish to thank my family and friends in China and Germany, as well as my in-laws in England for providing an incredibly resistant support system on which I could rely on. My parents in particular were always there to encourage me, assist me financially and offer practical sustenance in everyday life. With Ralf Schultheiss I discussed the topic of racism and identity in Germany over many hours of intense conversation, and I only regret that I was unable to do so with his daughter Pia, who was taken from us much too early. The birth of my son William Haichuan in 2013 right in the middle of my writing process brought home Hannah Arendt's phrase about the freedom and *Zauber* that lies inherent in all new beginnings. I thank especially my parents for helping in the challenging first years of toddlerhood, creating time for me to think and write, and to reclaim my identity as a political thinker.

Without my husband Lloyd West, this dissertation would not have been written and I would have felt too daunted to come to the U.S. for graduate study. We decided to embark on the life of the mind together back in Cambridge (when we barely knew what that meant) and have since created a traveling cocoon of home, love and liberation in our own hypermobile lives. Lloyd's incredible intellectual breadth ranging from pure mathematics, science, history, philosophy, Chinese art and literature, and dark British humor are a constant source of inspiration for my own work. He has served as first reader, editor, research assistant, coding software engineer and unyielding cheerleader for this dissertation. His anarchic mind gave me the courage to rethink the foundations of political theories about identity completely anew. What I owe to him existentially is beyond words.

This dissertation is dedicated to my parents, Heng Yu and Du Lan, and my late grandfather, 喻名武 . They are survivors of the worst kind of political dehumanization in the 20th century during the Maoist years. My grandfather, a highly gifted engineer who registered various patents, suffered unimaginable trauma during persecution and imprisonment. I hope that this dissertation can somehow show them that even the worst dehumanization can one day be partly transcended through one's children's efforts of trying to understand their own identity in the bewildering tides of history.

For my parents,
Yu Heng and Du Lan,
and my late grandfather,
喻名武

1. Introduction: Towards a "Political Brain Perspective"

> At present, neuroscience is a collection of facts still awaiting an over-arching theory.
>
> Gary Marcus and Jeremy Freeman, *The Future of the Brain*

> Nature hath made men so equall, in the faculties of body, and mind (…) from this equality of ability, ariseth equality of hope in the attainment of our Ends.
>
> Thomas Hobbes, *Leviathan*

> In any event, we must remember that it's not the blinded wrongdoers who are primarily responsible for the triumph of evil in the world, but the spiritually sighted servants of the good.
>
> Fyodor Stepun, *Foregone and Gone Forever*

1.1. So Much Brain Data, So Little Theory

In 2012, a research team at Oxford University consisting of people from experimental psychology, psychiatry and practical ethics wanted to find out whether the ingestion of a commonly used medication for treating heart conditions would lower racial bias in individuals (Terbeck et al., 2012). In their study, they asked White participants to ingest the beta-blocker medication 'propranolol' and tested for racial bias towards non-White faces. Propranolol had no strong effect on explicit racial bias, but when tested for subconsciously held biases through the 'Implicit Association Test' (IAT; Greenwald, McGhee, & Schwartz, 1998), propranolol appeared to significantly reduce participants' negative biases towards non-Whites. As expected, the media quickly absorbed these explosive findings through sensationalist headlines such as "A Pill That Could Prevent Racism?", "Blood Pressure Drug Reduces 'In-Built Racism'", and "Feeling Racist?

Blood Pressure Drug Propranolol May Open Hearts and Minds"[1]. One of the investigators of the study was Julian Savulescu, a leading philosopher of bioethics and director of Oxford's Center for Neuroethics, who had previously held controversial positions on selective embryonic procreation, infanticide, and euthanasia in the context of disability and comatic patients (Savulescu, 2006; Kahane and Savulescu, 2011; Persson and Savulescu)[2]. In the press statement accompanying the 2012 propranolol study, Savulescu however stressed that "propranolol is not a pill to cure racism", even though "such research raises the tantalising possibility that our unconscious racial attitudes could be modulated using drugs, a possibility that requires careful ethical analysis [because] biological research aiming to make people morally better has a dark history."[3].

Apart from the glaring ethical problems that the potential development and usage of an 'anti-racism pill' would entail, the 2012 study left some crucial questions unanswered: Did propranolol influence implicit racial bias directly or was the reduction of racial bias a side-effect of the pill's lowering of participants' heart rates? Did propranolol completely change racial attitudes or did it affect specific brain regions known to be implicated in implicit racial bias, and if so, which ones? In order to answer these questions, the 2012 research team conducted a follow-up study in 2015 that used functional magnetic resonance imaging (fMRI) to determine which brain regions exactly (if at all) were affected during exposure to non-White faces after ingestion of the drug (Terbeck et al.,

[1] http://www.telegraph.co.uk/news/health/news/9129029/Blood-pressure-drug-reduces-in-built-racism.html; http://www.nydailynews.com/life-style/health/study-shows-heart-disease-pill-lowers-racist-feelings-report-article-1.1034779; http://www.thedailybeast.com/articles/2012/03/13/can-the-heart-and-blood-pressure-pill-propranolol-cure-racism.html

[2] See also works by philosopher Peter Singer (1979), Savulescu's doctoral father, on similar controversial issues within medical ethics. For a sophisticated argument in favor of "cognitive enhancement", and the relationship between neuroscience and human rights, see Cohen (2015).

[3] http://www.neuroethics.ox.ac.uk/latest_news/propranolol_racism

2015). Back in 2012, the research team hypothesized that the reduction of implicit bias might have been due to propranolol's effect on the brain's amygdala activity – a brain region that had previously been shown to be impacted by propranolol (Talmi, Hurlemann, Patin, & Dolan, 2010), and that moreover had been connected to negative racial attitudes (Phelps et al., 2000; Hart et al., 2000) and emotional arousal in general.

However, the 2015 study yielded unexpected results: instead of acting on the amygdala, propranolol changed the way how initially threatening Black faces were processed as less threatening in the fusiform gyrus area, a brain area which also includes the fusiform face area (FFA), responsible for face recognition. In other words, the drug directly modified particular brain activity related to threatening face-perception and implicit racial bias. The fMRI results therefore confirmed a direct effect of propranolol on the neurobiological mechanisms underlying racial bias and ruled out the possibility that the phenomenon was a mere side effect of the drug. The pill and the brain were indeed connected[4].

The propranolol studies sum up everything that is intriguing and exciting, but also unnerving and troubling about current interdisciplinary research at the nexus of the cognitive and neurobiological sciences, and socio-political phenomena. Despite their dazzling results, what is striking about the propranolol studies is how little they can actually tell us what to do about racism at the political or societal level, to the point that one of the principal investigators warns that the findings do not imply that propranolol should be seen as a medical cure for racism – even though at first glance this is exactly what the results seem to imply and that is how the media interpreted them.

[4] For an in-depth, book-length discussion of these two studies and medical "cures" against prejudice see Terbeck (2016).

Although other studies on out-group exclusion, prejudice and racial bias in the relatively novel field of social neuroscience (Cacioppo, Visser, and Pickett, 2006; Todorov, Fiske, and Prentice, 2011; Schutt, Seidman, and Keshavan, 2015) and its even newer offspring, political neuroscience (Jost, Nam, Amodio, and Van Bavel, 2014), might not appear as striking or controversial compared to the propranolol studies, I argue that a similar uncertainty about the political, normative and moral implications is pervasive in many of these studies. Despite an ever-growing abundance of 'social brain data' that is being churned out by these disciplines – consider for example that in 2001, a Google search for 'social cognitive neuroscience' generated only 6 hits, compared to 52,000 hits in 2009 (Lieberman, 2010, p.144) and 3,160,000 hits in 2016 – the vast amount of this 'social brain data' remains undigested by the social sciences and the humanities, and is barely incorporated into political theorizing, quantitative models, policy debates or course syllabi at the research and teaching level in political science departments.

That being said, there exists within the social sciences a small but burgeoning subfield of promising interdisciplinary work on cognition and emotions in political processes (Redlawsk, 2006; Marcus, 2013) and international relations (Crawford, 2009), the biological conditions underlying politicians' decision-making (Wiegele 1985; McDermott, 2008), the evolutionary and genetic foundations of the political animal (Hatemi and McDermott, 2011), the relationship between political ideologies and cognitive mechanisms (Hibbing, Smith, and Alford, 2014; Huddy, Feldman, & Lowns 2014); neuroscience's influence on legal theory and new types of evidence presented in the court room (Pardo and Patterson, 2013), the socio-cultural implications of

neuroscience studies (Choudhury and Slaby, 2011), and the implications of neuroscientific revelations for political theory (Vander Valk, 2012) and political science more generally (Blank, Hines and Funke, 2014). This small subfield is driven by a growing number of political scientists, policy researchers and political philosophers who are intrigued by how brain data can reveal how our brains operate in social settings and interpersonal interactions. They realize that the brain data might be able to buttress or even confirm more speculative assumptions held by political scientists about the motivating reasons underlying human social perception and decision-making; likewise, new revelations about how our social brains function might also contradict hitherto entrenched beliefs about the nature of the political animal, serving therefore as a corrective tool with which political scientists can begin to reconceptualize and reframe existing political theories and models.

Despite the fact that this subfield is still in its infancy, and that therefore a comprehensive incorporation of 'social brain data' at the research and teaching level has yet to fully materialize within political science, a 'new biological turn' has slowly but steadily been taking place within the social sciences more broadly over the last three decades. Maurizio Meloni (2012) describes how various disciplines – including psychiatry, sociology, philosophy and political theory – have recently begun to overcome the strong anti-naturalist sentiment that emerged immediately after the Second World War, which back then was an initial reaction to the biological and evolutionary theories that had supported and justified horrendous atrocities such as the Holocaust, eugenic policies, and medical experiments carried out on the disabled and other vulnerable groups. The ensuing "mistrust and outright hostility" (Franks and Smith, 1999) that

defined the "relationship between the biological and the social in the twentieth century" (Meloni, 2012, p.27) was re-examined in the latter half of the century, starting with E.O. Wilson's *Sociobiology* (1975) and Richard Dawkin's *The Selfish Gene* (1976).

What followed were attempts at "re-alignment of the human social sciences with the life sciences" (Benton, 1991, p.25): through a revival of biological foundations for defining humanness (Richards, 1987; Degler, 1991; Laland and Brown, 2002), a revival of a biologically grounded concept of human nature (Arnhart, 1998; Masters, 1991; Konner, 2002), a new take on 'embodied political agency' in which the 'layered' dimensions of political thinking are acknowledged (Connolly, 2002; Thiele, 2006). Further to this there are more specific attempts at linking socioeconomic inequalities with white brain matter volume (Gianaros et al., 2013), linking cross-cultural mentalizing abilities with outcomes in International relations (Franklin, Stevenson, Ambady, & Adams, 2015), exploring the genetic basis and neurocognitive correlates for political orientations and party affiliations (Amodio et al. 2007; Alford and Hibbing, 2008; Alford, Smith and Hibbing, 2014)[5], and attempts by 'new feminism' to engage findings in genetics and neuroscience with a more materialist and embodied idea of female identity (Rippon, Jordan-Young, & Fine, 2014; Schmitz & Höppner, 2014), amongst other initiatives and trends.

Yet despite this 'biological turn' within the social sciences more broadly, there is much uncertainty (and little debate) about *concrete implications of neuroscience for core issues in Western political theory*, such as the social contract, deliberative democracy,

[5] For an excellent piece of criticism of *genopolitics*, attacking it both for its potential methodological and normative pitfalls, see Charney (2012).

human rights, multiculturalism, nationalism, or the politics of identity difference[6]. By this I mean for example specific insights about how studies on implicit bias against other-race faces would impact assumptions about inclusiveness held in multiculturalist theory, how the situational and precarious conditions under which rational decision-making takes place would affect social contract scenarios imagined by political theorists, or how our cognitive tendency to only fully humanize in-group rather than out-group members would alter strategies of persuasion in human rights education and dialogue.

In order to reach this level of concreteness, I suggest that political theorists who wish to draw meaningful conclusions and devise persuasive neuropolitical theories from the huge (and messy) array of brain data need to pay attention to the following four points:

1. Political theorists should engage in a critical and substantial examination of the **neuroscience methodology** employed in the brain studies that they wish to incorporate in their neuropolitical theories; asking, for example, what exactly an fMRI scan can and cannot tell us in regard to specific political questions and scenarios.
2. Brain data should not be treated as self-evident but be carefully **contextualized and adjusted** to the assumptions, ideas and language within the respective field of political theory to which political theorists wish to apply the data.
3. Political theorists should aim to develop a **new neuropolitical language** out of the brain data, especially in cases where certain brain insights seem to reverse or completely contradict the respective assumptions within a political theory. An awareness of the complexity of the actual political world into which this new neuropolitical language is potentially applied is crucial.
4. Finally, insights into our social brains and how they navigate the social world should be used in **expanding and adjusting larger philosophical ideas** about human nature, the political animal and the biopolitical foundations of the political self.

Despite some existing efforts, the extent to which 'social brain data' is seriously and substantively incorporated into political theorizing and policy-related debates in the

[6] For first attempts in this direction see Vander Valk (2012) on the relationship between political theory and neuroscience.

academic field is still very small; the manner in which this is carried out often lacks focus and suffers at times from conceptual sophistication and clarity. I contend that it is insufficient to broadly claim that 'emotions matter in politics' (McDermott, 2004), 'affect influences rational decision-making' (Marcus, Neuman, & MacKuen, 2000)[7] or indeed – in reference to the propranolol studies mentioned at the beginning – to jump to the conclusion that a racism pill could solve the causes of systemic racism. Meanwhile, it is too reductionist and sweeping to explain in-group membership and political activism primarily through the 'mirror neuron system' (Keestra, 2012; Newman-Norlund, Burch, and Becofsky, 2013), to ascribe voting behavior exclusively to certain brain mechanisms (Westen, 2008; Cory, 2004) or to establish a fuzzy link between the fluidity of globalized identities and brain plasticity (Malabou, 2008).

This dissertation is an attempt at contributing to this burgeoning field of critical and thoughtful political theorizing about our human brains. It aims to offer a detailed *neuropolitical theory about the effects of everyday dehumanized perception* in the context of a hyperdiverse body politic. It begins with the premise that no social brain research, no matter how insightful or groundbreaking, is self-evident in terms of its implications for political reasoning and application in the political world. As the propranolol studies showed, in order for neuropolitical theorizing to be persuasive, we cannot jump to conclusions by replacing political judgment with the brain data itself. If we want to avoid broad claims and over-simplified applications, the brain data needs to be carefully and painstakingly adjusted to the complexity of socio-political phenomena, especially when it involves normative implications.

[7] This is not to deny that there exist excellent interdisciplinary explorations about the implications of neuroscience research on emotions for the humanities, see for example Reddy (2001).

Moreover, no amount of social brain research can offer a complete source for political theorizing for us – the point here is that we are not looking for the brain data to fulfill an ontological purpose of defining who we are as political beings, but foremost help us understand how our brains function and what kind of underlying mechanisms are at play when we engage in politics with other human beings. This, of course, is grounded on an even more basic premise that when it comes to the study of politics, brain processes and mechanisms matter fundamentally in the first place. In order to forge tenable political theories as well as potential policy recommendations out of social brain research, I argue that we need to take all of the above into account.

This dissertation is aimed at *theory building* rather than trying to claim causal relationships between brain mechanisms and particular political behaviors. The kind of neuropolitical theory building that I attempt is *minimal*, in that I try to identify those cognitive mechanisms and brain capacities that I deem most threatening to the stability of the hyperdiverse body politic (and therefore most essential to identify and aiming to overcome), which I narrow down in this dissertation to one specific social brain capacity: *dehumanized perception of the other*. A detailed description of dehumanized perception from the social neuroscience viewpoint and the various manifestations of dehumanization in the form of animalistic, mechanistic, explicit and implicit dehumanization will follow shortly below – as well as an examination of the effects of dehumanizing others on one's own brain and the political effects of feeling dehumanized by others in turn.

1.2. The 'Political Brain Perspective': Why Brain Mechanisms Matter for Politics

In this section I want to develop what I call the *Political Brain Perspective* (PBP): an interdisciplinary perspective that allows us to analyze political phenomena through the lens of the social brain. I specifically call the PBP a *perspective* because I do not wish to make a causal or systematic point here; rather, what I wish to offer is a distinct perspective, thoughtfully and critically grown out of relevant brain data for the political question at hand, which allows us to access crucial brain mechanisms relevant for political analysis, and which enables us to theorize about the stability of political societies based on the cognitive abilities and limitations of its political actors. Adopting a PBP is therefore a vital first step before any neuropolitical theorizing can take place. Indeed, precursors of the PBP have been present throughout the history of Western and non-Western political thought, from Aristotle's preoccupation with psychology and biological taxonomy in *De Anima* and with the nature of human senses in *Parva Naturalia*, the Daoist philosopher Zhuang Zi's epistemological investigation into the fallacies of human perception and its consequences for constructing social reality; the dilemma facing the human mind and its passions in the context of social actions and going to war in the Ancient Hindu text *Bhagavad Gita*, Baruch Spinoza's protobiological reflections about affect and actions in the *Ethics*, Thomas Hobbes' concern in *Leviathan* and *De Cive* with the psychological conditions present in the state of nature and under the rule of a sovereign, and David Hume's search for empirically grounded empathetic sentiments in the *Enquiry Concerning the Principle of Morals*, to name a few.

In particular, Hobbes' employment of a PBP in trying to understand how to overcome the religious and social divisions of his time serves as an important guideline for this dissertation project. Hobbes' analysis of the senses, as well as the capacities and

fallacies of human imagination in Part I of *Leviathan* should be understood as a crucial strategic move: by deliberately placing his investigation of the cognitive, affective and physiological conditions of political 'subjects' *before* his discussion of the state of nature and the creation of a commonwealth based on peace in Part II, Hobbes makes clear that we cannot construct a political theory of the modern state – and for that matter, social cooperation and peace between divided social groups – without having an extensive understanding of the human mind's limitations and capacities. If Hobbes tried to build his political argument on a materialistic theory of the mind with the limited knowledge on human cognition available to him in the 17th century (Johnston, 1986; Tuck 1991), how would he have reacted to the proliferation of neuroscientific and psychological data in the 20th and 21st century? One can only speculate that Hobbes would have been intensely curious and eager to understand the groundbreaking insights into the workings of the social brain, whilst also being careful to apply these insights to the politics of the commonwealth in the same meticulously differentiated and innovative way as he did in *Leviathan*[8].

One of Hobbes' main aims in regard to overcoming certain cognitive conditions of his subjects was to free them from unnecessary fear. In Part IV of *Leviathan*, for example, he warns us of dangerous mental constructs stemming from misinterpretation of religious scriptures, such as superstitions and hallucinations, which can threaten the stability of the commonwealth. His underlying assumption was that the human mind was

[8] One might object here that Hobbes' was skeptical of the experimental method, as transpired in his exchange with Robert Boyle, see Shapin & Schaffer (1985). It is important to realize though that the 17th century idea of scientific experimentation differed considerably from today's brain imaging experiments, and that therefore, Hobbes' disagreement with Boyle does not necessarily mean that he would reject today's brain data insights.

naturally flawed in identifying threats in a rational manner, which for him had significant consequences for ensuring political stability between different religious and social groups over time. Most important, an "absence of any objective standards by which to measure what was right and wrong, or even what was beneficial or harmful to a human being" (Tuck, 1991, p. xxvi) led Hobbes to establish that the true nature of conflict was first and foremost "a conflict of belief" (ibid.). In *De Cive* he states:

> "For if the patterns of human action were known with the same certainty as the relations of magnitude in figures, then ambition and greed, whose power rests on the false opinions of the common people about right and wrong, would be disarmed, and the human race would enjoy such secure power that (apart from conflicts over space as the population grew) it seems unlikely that it would ever have to fight again."
> (*De Cive*, Epistle Dedicatory 6)

What transpires from this quote is Hobbes' recognition of the potentially *destructive psychological forces driving social human interaction* and thus the inherent fragility of social bonds. Moreover, he acknowledges that uncovering the cognitive and affective mechanisms behind these forces is a necessary first step in establishing a peaceful commonwealth. One could argue that Hobbes' search for certainty in regard to the destructive mechanisms underlying political action is a clear sign of skepticism towards more benevolent or optimistic accounts of 'human nature' (Slomp, 2000). Instead of venturing into the long-standing discussion of whether Hobbes was a pessimist of human nature, what is significant here for the PBP is that Hobbes, through his search for certainty regarding the destructive psychological forces driving social human interaction, is setting the bar for *where politics begins* at a very high point.

What do I mean by this? Hobbes understood that the stability of his utopian Leviathan as the "Mortal God" and "Artificial Man" rests on the psychological regulation and self-control of its individual members. The cognitive capacities of the individual

members, however, are not necessarily ideal when it comes to mutual contractarian cooperation under a shared sovereign: Hobbes outlines in detail in Part I of *Leviathan* how human vision, sense, speech and imagination are all flawed when it comes to making decisions free from fear and based on reason alone. Indeed, reason itself, unlike sense or memory, is "attained by industry" and not biological inheritance (Part I, Chapter 5), meaning that humans have to prudently learn how to regulate their imagination and passions – only then will they be able to leave the state of nature, and politics in the form of peaceful cooperation can begin.

Hobbes' search for more objective cognitive standards of predicting when and how human beings will form stable and peaceful political unions is taken up again by political scientists in the 20th century, indirectly first by the behavioralist movement. Although I stress in the subsequent chapter that the cognitive revolution of the 1950s was a deliberate departure and stood in clear opposition to the 'behaviorism' school that preceded it, it is worth noting some shared goals in terms of theorizing political actions.

David Easton's piece on *The Decline of Modern Political Theory* (1951) is significant here: Easton, who bore major responsibility for introducing the behavioralist sciences and systems theory into political science in the mid-20th century, argued that current political theory had failed to take on the "task of building systematic theory about political behavior". Easton's search for recurring patterns and a more comprehensive and generalizable theory building surrounding the capacities and behaviors of political actors resonates with Hobbes' own endeavors. In the U.S.'s mid-20th century political science environment (particularly at the University of Chicago), WWII European émigrés such as Leo Strauss (1953) and Eric Voegelin (2000[1952]) lead the attack on Enlightenment

optimism about scientism and the idea of historical progress. Easton cautiously reacted against this with his search for a new empiricism for political science, which despite its focus on scientific methods and models was driven by profound normative concerns (Gunnell, 2013, p.197). Even though behavioralists such as Easton were not particularly interested in accessing political processes at the brain level, it is worth noting his shared desire with one of the PBP forerunners, Thomas Hobbes. Both tried to understand more deeply embedded patterns and universal mechanisms underneath the messy reality of political phenomena.

In this sense, developing the PBP for political theory analysis within this dissertation is only a natural extension and deliberate transformation of previous attempts within political science (Lasswell, 1936; Merriam, 1945; Almond & Verba, 1963; Easton, 1965) to uncover meaningful and crucial determinants of political perception and decision-making, and understand which inherent human abilities undermine the stability of a political community. In this dissertation, the PBP serves the purpose of establishing a *minimal standard for when politics begins*, in the Hobbesian sense of politics as peace and security from mutual destruction, but also as the ability of individuals to recognize each other as equally human and cooperate with each other on this crucial basis. The PBP differs from other, non-cognitive perspectives in that it puts the *cognitive abilities and limitations of political actors* at the core of any political analysis.

There is a reason why the PBP is not called the 'Political Mind Perspective': the PBP recognizes the physical human brain as the center of the political individual's emotions, perception, attitudes, beliefs and behavior, in so far as that the neurobiological manifestations of emotions, perception, attitudes, beliefs and behavior – as measured

through functional magnetic resonance imaging (fMRI), electroencephalography (EEG), transcranial magnetic stimulation (TMS), brain lesion studies and other methods – are treated as significant and fundamental for understanding some aspects of political phenomena. However, the PBP is *not* reductionist in that it does not claim to understand completely or explain exclusively all aspects of political phenomena through neurobiological processes, but merely pointing out one (albeit fundamental) layer of bricks within a complex house structure.

In the philosophy of mind, dualists believe in the distinction between mind and brain (or mind and body) because for them, these are two radically separate entities. Dualism postulates that the mind is independent from the reality of the physical world (including the physical human brain), in that mental states such as thoughts, beliefs, desires and imagings exist separately from any neurobiological processes (Descartes, 1996 [1641]; Swinburne, 1986; Robinson, 1982, 2016; Hodgson, 1988; Herbert, 1998). By placing the 'brain' instead of the 'mind' at the center of its perspective, the PBP is making a decidedly *non-dualist statement* here but not an anti-dualist one: the PBP does believe that neurobiological processes matter and are linked to mental experiences but the PBP does *not* insist that all mental concepts can be mapped in their entirety at the brain level.

The PBP does not intend to weigh in on the long-standing debate within the philosophy of mind between 'intentional realists' (Fodor, 1987; Dretske, 1988) and 'intentional eliminativists' (Churchland, 1981; Dennett, 1987), on whether commonsense psychology (or folk psychology) is able to capture a level of reality about our mental lives that modern scientific theories are unable to access. This debate is linked to the

general debate within philosophy of mind on mental representations, and the famous challenge that philosopher Thomas Nagel (1974) posed to cognitive science, in claiming that the latter will never uncover what it feels like to be inside the mind of a bat. I will return in chapter 2 to this epistemological challenge on how cognitive science is unable to fully access phenomenological experiences of mental processes, and to which extent this is politically relevant. I wish to stress here that the PBP is *non-reductionist* in that it does not claim to capture all aspects of our mental experiences and processes – even though it does believe that including fundamental brain processes in the political analysis of specific political phenomena (especially those that involve interpersonal social perceptions and dynamics) is necessary in order to understand an important part of political reality.

In addition, the PBP takes a *non-determinist stance* towards the genetic causation of brain processes and human behavior. Edmund T. Rolls, the director of the Oxford Center for Computational Neuroscience, discovered that only about 15% of our genes determine brain connectivity (i.e. roughly 2,000-5,000 genes out of a total of 30,000 genes that each human being carries) (Rolls and Stringer 2000), which means that "genes can only specify some of the general rules of brain wiring" and that therefore the "connectivity of the brain must be specified by self-organizing processes including learning from the environment" (Rolls, 2012, p.5). In other words, much of the way in which our social brains are structured and function is determined by how they adapt to particular environments and circumstances.

From an evolutionary viewpoint, any genes influencing specific goals ought not affect behavior too strongly because animals and humans have to adapt the goal of their

reward according to environmental availability (e.g. food, water, shelter, sex); thus the determinist effect of genes should not be overestimated. Perhaps the most crucial way in which genes influence our brains is that they "adapt our minds to make certain stimuli, events, and interactions rewarding or punishing" (ibid., p. 4). In addition, there exist individual differences in humans' sensitivity to various rewards and punishments. At the brain level, this translates as 'noise' in the form of the randomness related to neuronal firing times (Rolls and Deco, 2010).

In the context of social and political interactions, this can include the rewarding aspects of in-group belonging (Golby, Gabrieli, Chiao, & Eberhardt, 2001; Isobe, Nakashima, & Ura, 2013), out-group pain (Cikara, Bruneau, Van Bavel, & Saxe, 2014), partisanship (Knutson et al., 2006) and the winning of one's own in-group team (Bornstein, Sagiv, & Halevy, 2008), as well as the punishing experience of being ostracized from one's in-group (Zadro, Williams, & Richardson, 2004), and feeling victimized (Leidner, Castano, & Ginges, 2013) and dehumanized (Kteily, Bruneau, & Hodson, 2016) by others. Genes are unable to determine towards which stimuli exactly our brains will react to, but they can predispose us to seek certain rewarding social experiences (and avoid punishing ones) that solidify our sense of in-group belonging. This can manifest politically for example in nationalism (Carter, Ferguson, & Hassin, 2011) or partisanship and party affiliation, but with variation within the population.

At its core, the PBP is an attempt to *bridge the chasm between the physical reality of the brain and the phenomenological experience of political conflict* (and cooperation) between different groups of individuals. The PBP is necessary because the empirical data depicting physical brain processes is not self-evident, therefore requiring a distinct

political language to translate brain insights into political concepts, which can then be applied to the realm of politics in a convincing and justifiable way. This political language is necessary and even indispensable because the brain data itself (as presented in social neuroscience and psychology papers) cannot be applied directly to political problems – as we have seen in the propranolol study – without the help of a conceptual and linguistic 'bridge' (i.e. the PBP). The purpose of this 'bridge' is to articulate the pragmatic, analytic and normative implications that might arise when the brain data is applied to complex political problems, and offer an interdisciplinary framework in which these implications can be explored.

1.3. Towards a Minimal Neuropolitical Theory

The aim of this dissertation is to sketch out the beginnings of a neuropolitical theory for hyperdiverse societies. For this we need to identify and understand the fundamental cognitive mechanisms that can threaten the peaceful co-existence and political cooperation between different social groups. The first question has to be: '*What is the political question that we are bringing to the brain science*?' The conscious formulating of an interdisciplinary political question is vital for laying the foundation of a minimal theory. The point here is *not* to simply treat the brain science as an exciting new tool that is then applied to a political problem picked from a traditional list of long-standing issues within political science (such as voting behavior, decision-making, ethnic conflict etc.). The way to integrate the brain data should not consist in merely treating it as a novel independent variable for quantitative political scientists, or as empirical evidence for confirming political commitments and philosophical beliefs of political theorists.

By applying a neuropolitical perspective to politics, I argue that we are fundamentally changing the way we conceptualize politics itself – indeed, this is inevitable if we want to formulate sensible political questions that can absorb the brain data in a meaningful way. In the social sciences (as in all sciences) researchers always rely on certain paradigms – about human cognition and behavior, for example – to construct their hypotheses about the world. By not taking for granted the paradigms underlying long-standing political problems, I propose a different kind of neuropolitical theorizing where we are willing to ask anew what are the cognitive conditions of individuals entering politics and what is the threshold of when politics can begin. This kind of approach might change what kind of political questions we are asking and possibly how we conceptualize politics itself.

Thomas S. Kuhn, who famously examined the inherent dynamics of paradigm changes within natural sciences, wrote in *The Structure of Scientific Revolutions* (1970 [1962]) about exactly this issue, stating that each scientific revolution

> "produced a consequent shift in the *problems available for scientific scrutiny* and in the standards by which the profession determined what should count as an admissible problem (…) each transformed the scientific imagination in ways that we should ultimately need to describe as a transformation of the world within which scientific work was done [emphasis added]" (p. 6).

According to Kuhn, paradigm shifts – of which the PBP aspiring to be one – lead not only to changes in people's previous world view but most important, to significant changes in the kind of questions and problems we want to answer and tackle. The question, therefore, of 'What is the political question that we are bringing to the brain science?' is an acknowledgement of the necessity to readjust and revisit more

fundamental ways in which we conceive of political phenomena, and what the puzzle within that phenomenon is.

Kuhn warned that "in the absence of a paradigm (…) all of the facts that could possibly pertain to the development of a given science are likely to seem equally relevant" (ibid., p.15). This is the current state a majority of interdisciplinary theorizing on the political brain has found itself to be in – by neither fully committing to previous paradigms about human cognition nor to a new paradigm of the political brain, all brain facts seem equally relevant.

For example, this becomes evident in the field of the study of emotions and politics: on the question of emotions' influence on rational decision-making, if the political question that we bring to the data is not consciously and specifically formulated to operate within an interdisciplinary framework, then any kind of genetic or neurobiological data on emotions can appear as equally significant. The point here is not to simply replace our old assumptions about the role of emotions with the new brain data, but to reformulate our most fundamental political questions by taking into account the brain data and the complexity of the political problem in equal measure.

In this vein, the PBP tries to formulate a distinct language for theorizing about politics, primarily by looking at how the language of brain data (and the language used by neuroscientists to explain this data) maps onto the existing language within political theory. The aim is not to just replace one language with the other but to formulate concepts and a mode of debating about these concepts that is located exactly at the interdisciplinary nexus of neuropolitics. I argue that this is in the interest of cognitive scientists because many normative questions that arise from the brain data cannot be

answered through the linguistic and conceptual tools of social cognitive neuroscience itself; meanwhile, this is also in the interest of political theory, since even if political theory wanted to acknowledge the existence of brain mechanisms in politics, many insights into these cognitive mechanisms would not be empirically and linguistically available to it without the inclusion of PBP. In other words, both disciplines need each other in creating a currently absent space where the brain data is digested, critically reflected upon, and transformed into a theory and language that can be applied to debates within political theory and political science more widely.

1.4. Vulnerable Brains: Why Dehumanization Matters for Hyperdiverse Societies

The neuropolitical question that this dissertation explores is '*What are the minimal cognitive conditions for living together in a hyperdiverse body politic?*' In particular, the focus of this dissertation is to develop a PBP based on the recent psychological and neurobiological evidence on *dehumanized perception* of out-groups (Abelson, Kinders, Peters, & Fiske, 1982; Leyens et al., 2001; Mitchell, Macrae, & Banaji, 2005; Haslam, 2006; Harris & Fiske, 2006, 2007; Hodson & Costello, 2007; Goff, Eberhardt, Williams, & Jackson, 2008; Vaes & Paladino, 2010; Bain, Vaes, & Leyens, 2014; Waytz, Hoffman, & Trawalter, 2013; Kteily, Waytz, Bruneau, & Cotterill, 2015), and to discuss what effects this kind of dehumanization has on politics at the domestic and international level.

The picture that is emerging on dehumanization and its relation to social cognition is that dehumanizing other humans is an everyday, often subtle phenomenon (Leyens et al., 2001) that we all engage in as part of how we function socially (Cameron, Harris, & Payne, 2016). Although dehumanization has been studied within social psychology at the

behavioral for over a couple of decades, exploring the neural underpinnings of dehumanization is a relatively novel endeavor. What is significant here is that human beings can switch extremely quickly and often effortlessly between humanizing an in-group member or even anthropomorphizing non-human objects around them, and dehumanizing members from other social, cultural and political groups (Harris & Fiske, 2009). In *Invisible Mind: Flexible Social Cognition and Dehumanization* (2017), Lasana T. Harris treats the automatic and spontaneous ability of human beings to deny humanness to others as part of what he calls our 'flexible social cognition' system. 'Social cognition' is understood here as our ability to imagine and infer the mental states (mental state inference) of other individuals, thus one fundamental, neural aspect of dehumanization is the inability to infer someone's else's mental state. I will discuss in the subsequent chapters how this ability to infer someone else's mind is part of a wider, more established field that studies mentalizing and 'Theory of Mind' (Frith & Frith, 2005), and is further understood to be the building foundation for more complex emotions such as empathy and compassion (Singer & Klimecki, 2014).

Socially and politically – based on the insight into these neurocognitive mechanisms – dehumanization can be viewed as facilitating disastrous events such as intergroup aggression, torture and mass atrocities (Castano & Giner-Sorolla, 2006; Harris & Fiske, 2011), or have more subtle political effects, such as in the form of neglect of vulnerable out-groups ([SCM behavioral study]), rejection of refugees (Dalsklev & Kunst, 2014), support for stronger retributive punishment in the legal context (Capestany & Harris, 2014) and hostility towards social welfare programs (Huddy, Feldman, & Lowns 2014).

As a result, I treat the evidence on dehumanization as a highly significant kind of out-group categorization and exclusion, and therefore as a pivotal game changer for political theory. I argue that the recent psychological and neuroscience data on dehumanization is highly relevant for theorizing about politics in hyperdiverse societies, to the point that any minimal theory of social cooperation needs to be aware of it. Unlike other, more novel social neuroscience topics that are almost completely dependent on the recent availability of brain scanning and mapping methods, such as the 'mirror neuron theory' (Rizzolatti & Craighero, 2004; Iacoboni et al., 1999; Ramachandran, 2011) or the concept of 'neuroplasticity' (Maguire et al., 2000; Park & Reuter-Lorenz, 2009), dehumanization has been a long-standing topic within social and political psychology for several decades (Bandura, Underwood, & Fromson, 1975; Bar-Tal, 1989; Staub, 1989; Kelman, 1973; Opotow, 1990; Mikulincer & Shaver, 2012). Although the theory of dehumanization is not devoid of its own methodological and normative controversies, one benefit of employing this theory as the basis for a minimal neuropolitical theory is that it is still far less controversial and possibly more tenable than other, more novel, brain-based theories of social cognition.

This being said, what can a neuropolitical theory of dehumanization offer to the existing political discourse on prejudice, exclusion and racism, for example? Why do we need to establish a brain-based foundation for making sense of socio-political phenomena such as racial exclusion? One reason for this is that a neuropolitical theory of exclusion can add the dimension of a *tangible, physical and neurobiological reality* to the phenomenon of identity exclusion. Why does this matter?

W.E.B. DuBois, the great American sociologist and philosopher of race, insisted at the turn of the 20th century (and at the height of the New Imperialism) in his *The Conservation of Races* (1897) that race had no scientific foundation, yet he also believed that the experience of race and racism was nonetheless very real. Du Bois rejects the practice of delineating race difference along "color, hair, cranial measurements and language", pointing out how all of these features have been intermingled amongst social groups across human history. Yet Du Bois is not willing to give up the concept of race, chiefly for political reasons. For DuBois, the yet unaccomplished attainment of equality for oppressed Black people living in the U.S. is a clear indication that the division of political power and social resources along racial lines is still alive and very real. It is for this reason that "Negro people [have to] (…) maintain their racial identity until [these questions of policy and right] (…) and the ideal of human brotherhood has become a practical possibility".

Coming from a different point of analysis but indirectly making a similar point, Jean-Paul Sartre, in his piece *Anti-Semite and Jew* (1948) describes (or caricatures) the liberal "democrat" of his time for whom no racial categories exist but only the universal idea of mankind. Sartre rejects the "democrat's" denial of racial and social group difference for similar reasons to Du Bois because he contends that in order for Jewish people to be truly liberated, a "Jewish consciousness" and group identity is a necessary first step to achieve this; the "democrat's" viewpoint simply de-politicizes and denies the very real existence of exclusion based on group identity and difference. Similarly, Hannah Arendt, in an interview from 1964 in which she reflects on her escape from Nazi Germany and political activism on behalf of Jews during the WWII, says "If one is attacked as a Jew, one must

defend oneself as a Jew. Not as a German, not as an upholder of the Rights of Man, or whatever." (Gaus, 28 Oct 1964).

Du Bois, Sartre and Arendt are all trying to make a similar point within the discourse on exclusion and identity: none of them supports the oppression of one identity group in society by another, yet all of them stress the need to acknowledge the very real occurrence of the denial of humanness of one group by another, as well as the visceral experience of racial exclusion experienced by both the excluder and the excluded. Stating this aspect of social reality might be painful, which is perhaps why Sartre's "democrat" tries to deny it, but necessary in order for systemic and genuine political change to take place.

A neuropolitical theory of exclusion can strengthen this particular point Du Bois, Sartre and Arendt are trying to make: by helping to *establish the 'realness' of social exclusion and dehumanization* through locating exclusion at both the neurological level of the excluder and the experience of the excluded. A neuropolitical theory of exclusion, in other words, can add a biologically grounded aspect of reality to the multifaceted phenomenon of identity exclusion, without aiming to reduce or deny the interpretative, phenomenological and subjective complexity or structural inequity of the phenomenon itself. This kind of theory in form of the PBP can help buttress claims that racism, sexism, colonialism, dehumanization and other forms of oppressive exclusion are in fact very real – from the viewpoint of our brains and how our universally shared excluding brains affect social cognition, emotions and behavior. The PBP can become an empowering tool for excluded groups to establish the neurobiological manifestation of

exclusion and prejudice as a tangible reality, and from this vantage point, to make a more materialist and empowered argument about recognition and equality.

By locating a significant explanatory basis of exclusion in our brains – in the case of dehumanized perception, in the brain of the excluder – we might also be able to destigmatize, on a biological level, the human tendency for bias and flexible dehumanization of others. This does not mean that we have to accept this biological fact without any boundaries and conditions in the political realm (i.e. many behavioral or attitudinal manifestations of dehumanization should be rejected politically), or that we fatalistically surrender to our neurobiological preconditions of bias and exclusion. On the contrary, only by going through the painful realization that the vulnerable, dehumanizing brain is a fundamental part of all social humans and that none of us is morally or biologically superior in terms of our cognitive preconditions, can we actually begin to effectively tackle political and social instances of exclusion and the destabilizing effects everyday dehumanization can have on a political polity[9].

Related to this, Frantz Fanon in *White Skin, Black Masks* (1967) argues that colonialism's influence on the colonized person's sense of self and her internalized sense of inferiority should not chiefly be pinned down to the colonized person's individual neurosis – as a classic psychoanalytic analysis in the line of Freud, Jung, Adler and Lacan would attempt – but to the general social condition and the colonizer. The PBP thus allows an externalization of the experience of political and social exclusion, by making sense of the cognitive mechanisms and abilities of the excluder's brain, yet treats these

[9] Social Psychological Answers to Real-Word Questions "do-tank" (SPARQ) at Stanford University (https://sparq.stanford.edu/).

excluding and dehumanizing capacities as an integral part of who we are as political beings.

The uncomfortable fact is that out-group exclusion is not just perpetrated by the selective few but that dehumanized perception is integral to how we navigate ourselves socially and politically. Svetlana Aleksievich, in her oral history book *Secondhand Time: The Last of the Soviets* (2016), in which she tries to capture the complex history of the Soviet Union's rise and endurance through the stories of historical witnesses and participants, particularly by drawing out the conflicted ambiguities of human behavior in the face of revolutionary fervor, totalitarian submission and interpersonal relationships. She cites Fyodor Stepun, himself a survivor and witness of totalitarianism, who warns that the causes for evil in the world are often not just a few, misguided perpetrators but potentially also those who contend that their political actions are morally infallible and superior due to the noble nature of their beliefs – again, Sartre's "democrat" who believes to have transcended the phenomenon of racial exclusion through his abstract commitment to universal values comes to mind.

Here it becomes important to point out that dehumanized perception can also be necessary or even beneficial in certain social settings, such as in the medical sector when surgeons are able to dehumanize the patients they are operating on, in order to carry out their tasks most efficiently and with complete focus on the operating procedures, without the involvement of their personal emotions of empathy (Haque & Watyz, 2012; for a comprehensive review on this see Leyens, 2014). This goes again to show that the mere brain data on our flexible, dehumanizing minds is insufficient in its ability to offer direct political answers and frameworks for our most pressing questions concerning the

recognition, cooperation and solidarity between different groups in hyperdiverse societies. Our vulnerable, dehumanizing minds are neither morally or politically 'evil' or 'bad' per se, but they have the potential to disturb and destabilize various political projects and social relationships. The task of this dissertation is to identify which these are, how exactly the brain data would apply in the form of the PBP, and how our awareness and concrete measures can aid to overcome some manifestations of dehumanization that are politically unacceptable.

1.5. Chapter Overview

Chapter 2 offers a political history of the 'cognitive revolution', which contributed to the birth of social neuroscience, and subsequently, political neuroscience. The chapter argues that two key historical events – WWII and the Civil Rights Movement – were crucial in spurring the need to redefine 'human rationality' in socio-political contexts, as well as to find alternative methods to access implicit racial bias and prejudice. The purpose of providing this historical backdrop is to contextualize the PBP in the intellectual, multidisciplinary history out of which social and political neuroscience grew, showing its relevance for major political theory questions asked at the same time in 20th century history. Showing how a sophisticated and complex self-critical discourse developed *within* social cognitive science about neuroscience epistemology and methodology also helps to humanize cognitive scientists to those political theorists who fear that they are dealing with a crude reductionist and determinist opponent[10]. Chapter 2

[10] See Fiske & Dupree (2014) for a study on how non-scientists often stereotype scientists as non-trustworthy, even though they might be viewed as competent.

combines historical contextualization with a 'political philosophy of science' analysis of the current brain imaging techniques used in experimental neuroscience research.

The main argument there is that the brain data cannot tell political theorists *who* we are as political beings in terms of ontology, but rather, allows us to make claims about what our brains *do* when they engage in politics in terms of their cognitive mechanisms. The PBP therefore favors a 'brain mechanism' over a 'brain ontology' framework; which is subsequently applied in Chapter 4 when redefining the concept of "human dignity" in human rights and multiculturalism debates. Exposing the methodological possibilities and limitations behind the brain sciences for explaining political phenomena, as well as taking a clear epistemological stance in favor of a brain mechanism-based approach are both necessary for the arguments in the subsequent chapters and the wider objective of this dissertation to convincingly hold together and make sense to the reader.

In Chapters 3 to 5, I apply the PBP to various political theories and real-world political scenarios in which an insight into the brain capabilities for humanizing other groups and individuals is vital. This dissertation focuses on the political relevance of dehumanized perception at the brain (and partially behavioral) level. Chapter 3 sets out to define some basic neurocognitive responsibilities public officials such as elected politicians, judges and law enforcement officers ought to fulfill in order to justifiably and effectively represent diverse constituencies in liberal democracies. One of these neurocognitive responsibilities is defined as *mentalizing*, which plays a crucial role in perspective taking, humanization and empathy for others. Furthermore, Chapter 3 studies the role of *mentalizing* for public representatives in *two different frameworks of politics* in the liberal democratic state, where the actual cognitive time allotted for public

representatives to overcome potential dehumanization of constituents plays a decisive role. From a more theoretical viewpoint, Chapter 3 examines how the everyday and flexible aspect of dehumanized perception fits into previous political theories about human rationality, arguing that both liberal and Marxist thought have overlooked dehumanized perception as a fundamental way of how our social brains try to navigate themselves in the political world. This in turn has implications for theories about the politics of recognition, multiculturalism, and social contract theories.

Chapter 3 is intentionally broad in its scope, trying to show the relevance of dehumanization through the PBP for a multitude of political issues within various liberal democratic theories about cooperation, recognition and solidarity in divided and hyperdiverse societies, as well as for those public representatives who operate in these societies.

Chapter 4, in contrast, dissects one topic in particular, namely that of the idea of 'human dignity' in liberal universalist human rights theories and within current multiculturalism debates. Chapter 4 returns to the brain ontology vs. mechanisms debate that Chapter 2 touched upon, working through it in greater detail and with a clear focus on the 'dignity' concept. In a nutshell, Chapter 4 argues that 'human dignity' should not be construed as an ontological essence but as an interpersonal, neurocognitive ascriptive process where dignity can be ascribed flexibly in terms of (de)humanized perception of the other. Furthermore, by bringing in evidence from cultural neuroscience, the chapter argues that the effortless switching between different cultural identities *within* bicultural individuals shows that the current monocultural, predominantly Western individualist idea of 'dignity' needs to be updated for culturally hyperdiverse societies.

Finally, Chapter 5 applies the dehumanization argument to IR theories of civilizationary progress and clashes in world politics. This chapter looks at the political effects on intergroup conflict when a group feels dehumanized by another, arguing that current "civilizationary clash" theories that pit "civilized" nations against "barbaric" ones fuel the dehumanization of so-called barbaric groups and potentially encourage retaliatory and violent reciprocal dehumanization by those very groups. The chapter illustrates its point by examining various case studies, such as the use of the barbaric concept in Antiquity, the Israeli-Palestinian conflict, the Bosnian genocide, and the civilizationary rhetoric used by present-day right-wing parties in Western Europe.

Chapter 6 offers brief conclusionary remarks about the epistemological, empirical and normative insights gained in having applied the dehumanization argument through the PBP in this dissertation. It concludes by reflecting about what an understanding of the social human brain in terms of its flexible and innate ability to humanize others can offer us when devising political theories about cooperation in hyperdiverse societies, with a particular focus on the challenges for cross-cleavage political solidarity between groups.

2. From Political Mind to Political Brain: A Historical and Methodological Discussion

2.1. Adding Brain-Based Methods: What is There to Gain (and Lose)?

This chapter outlines how social neuroscience methods and techniques matter for long-standing questions in political science, as well as for the more specific aim of this dissertation: building a minimal neuropolitical theory for cooperation in hyperdiverse societies. However, brain imaging methods and techniques did not suddenly appear out of a vacuum space; they were created and contested through contentious and contingent historical trajectories across different disciplines, and continue to be debated and refined within the current fields of social and political neuroscience.

Rather than making a tight causal argument, the purpose of this chapter is to offer a *discursive* account of the political history of the 'cognitive revolution' and how it impacted epistemological and methodological outlooks in multiple fields. This chapter aims to draw attention to the complex social and political driving factors behind the development of brain imaging techniques in social neuroscience, and show how intimately they are connected to similar endeavors in the political theory field in the 20th century and beyond. Whereas Chapter 1 presented an analytic defense of the PBP, the point of this chapter is to defend the PBP through *contextualizing* it in the intellectual history of both cognitive science and political theory – both kinds of approaches are necessary for a comprehensive theory building of the PBP.

Most important, this chapter engages in a detailed discussion about methodological opportunities and pitfalls inherent in fMRI and other kinds of brain imaging techniques, particularly when it comes to application to political problems. The level of detail and seriousness dedicated to the technical aspects of brain imaging methodology in this chapter is necessary in order for the application of the PBP to more concrete political problems in the subsequent Chapters 3 to 5 to make sense.

At least since the middle of the 20th century, political science has borrowed methods and concepts from psychology to explain various political phenomena, ranging from the analysis of authoritarianism and political personalities, political ideology, voting choices and media influence on the electorate, to revolutions and rebellions, interpersonal conflict and violence, and ethnic and cultural identity politics (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Tetlock, 1983; Simon, 1985; Long, 1981; Lau & Sears, 1986; Elster, 1993; Green & Shapiro, 1994; Winter, 2000; Deutsch & Kinnvall, 2002; Jost & Sidanius, 2004; Huddy, Sears, & Levy, 2013). John Jost and Jim Sidanius posit that political psychology has become a "dynamic subfield that addresses the ways in which *political institutions* both affect and are affected by *human behavior* [original emphasis]" (Jost & Sidanius, 2004, p.1). This is echoed by another recent definition, which states that

> "At its core, political psychology concerns the behavior of individuals within a specific political system. Psychology alone cannot explain the Holocaust, intractable conflicts, war, or most other behavior of states or collective political actors in complex environments. Individuals do not act within a vacuum. Their behavior varies with, and responds to, differences in political institutions, political cultures, leadership styles and social norms."
>
> (Huddy, Sears, & Levy, 2013, p.3)

Both of these definitions highlight how political psychology is primarily concerned with the behavioral aspect of politics, not so much the cognitive and neurobiological part. This could be credited to more circumstantial factors, such as the lack of availability of brain-based methodologies when political psychology first took off mid-20th century, but possibly irrespective of historical circumstance, the focus on 'behavior' might be a conscious choice for a majority of political psychologists up until today. On various levels, this makes sense: an individual's behavior is the most obvious and tangible manifestation of the inner workings of their mind, and has the most direct consequences for political interactions and transactions.

However, behavior is only one component on a scale of 'mind manifestations', which can range from behavior, opinion, belief, attitude, affect, social cognitive processes and implicit bias, to the interplay of hormones and genetic factors, amongst many others. The aim of this chapter is to explain what the *Political Brain Perspective* (PBP) and its focus on social cognition at the brain level can contribute to political analysis, in addition to what more traditional psychological methods (such as behavioral and survey-based measures) offer political scientists already. I argue that the insights derived from using functional magnetic resonance imaging (fMRI), electroencephalography/event-related-potentials (EEG/ERPs) and brain lesion studies can be helpful in illuminating particular aspects of political phenomena that might otherwise be missed or ignored.

This chapter examines the benefits, drawbacks and reliability of these methods for political science, both from an epistemological and normative perspective. In particular, I assess to which extent the brain-based insights into implicit bias, dehumanization and 'flexible social cognition' (Harris, 2017) can contribute to a *minimal*

neuropolitical theory of cooperation and wellbeing for hyperdiverse societies. What can an fMRI scan about the denial of humanness contribute to existing political theories about intergroup conflict, recognition, and dignity? How do EEG studies on the rapid and automatic nature of implicit racial biases affect the way we conceptualize the political salience of time (e.g. exposure and contact of police with racial minority groups, time framework within which different racial groups interact regularly, etc.) during conflict? Further, how do fMRI and EEG/ERP techniques that gauge neuronal electric activity and blood flow changes at the brain-level alter the way we conceptualize social cognition and political perception, compared to behavioral and attitudinal methods more commonly used by political psychologists and quantitative political scientists? In the same vein, how do these neuroscience techniques alter the way we conceptualize the political self and the hyperdiverse body politic, compared to theoretical speculations and normative theorizing usually carried out by political theorists? How do these different levels of analysis of the same empirical political phenomenon or theoretical problem compare and contrast with each other, and can they possibly be complementary?

Based on the PBP, I firstly argue within this dissertation that brain data should *not* be treated as information about the ontology of the political self, meaning that quantitative brain data (no matter how much we accumulate now or in the future) cannot tell us who we *are* as political beings. Instead, brain data tells us what our brains *do* when engaged in certain social and political processes and in response to specific stimuli, revealing the neurobiological and cognitive *mechanisms* involved. Methodologically, this is reflected in how this dissertation employs brain data to compare it with assumptions made in existing political theories and to draw normative conclusions: the aim here is not

to delineate the ontology or 'essence' of the political self (or political mind), but to use the insights into social brain function to construct a plausible minimal theory of cooperation and wellbeing in hyperdiverse societies.

Secondly, recognizing that there exist universally shared brain functions amongst humans can become a compelling basis from which to establish which cognitive responsibilities different political actors owe to each other, irrespective of differences in status and group membership. For example, do different political actors need to humanize each other at the brain level before they can enter the social contract, and if they do not, how does this affect the stability of the social contract over time? Recognizing that there exist universally shared brain functions amongst humans also allows us to make more informed distinctions between *who exactly needs to be in command of their cognitive (de)humanization abilities the most*: whereas ideally, it would be desirable if all citizens and public representatives would be equally able to command and control their (de)humanization abilities, the realities of power inequalities (both in terms of sheer force and political power inequalities) between public representatives such as elected politicians, law enforcement officers, judges, etc. and ordinary members of the political community such as private citizens, undocumented immigrants, migrants and refugees suggest that the main burden, politically speaking, of avoiding dehumanization and encouraging humanization of others within one's brain lies with public representatives rather than the ordinary members of the political community (even though the latter have responsibilities as well). I expand on this discussion in Chapter 3.

The point of this chapter is to examine how we can justify the usage of methods from experimental neuroscience to make political arguments, to understand where and

how the 'cognitive worldview' emerged historically, how it fits in with more traditional methodologies hitherto employed in political science, and finally, how the PBP – which this dissertation is developing conceptually – can take advantage of the benefits that cognitive neuroscience methods offer, whilst also trying to avoid its inherent drawbacks and pitfalls. This chapter will defend in particular social neuroscientists' John Cacioppo and Penny Visser's (2003) commitment to a "multilevel integrative analysis", which aims to acknowledge the various levels of analysis – of which the neural level is only one of many – that contribute to a complete understanding of social cognition. "Multilevel integrative analysis" is a particularly helpful framework for addressing (and potentially overcoming) methodological controversies inherent in neuroscience research, such as reductionism and determinism, and for uniting a variety of cross-disciplinary explanations of social cognition under a wider research objective.

2.2. Back to the Beginnings: From the 'Cognitive Revolution' to Social Brains

George A. Miller, one of the founders of the 'cognitive revolution' that led to the emergence of what we today consider the field of cognitive neuroscience, dates its conception back to the specific date of September 11, 1956, where a symposium held at the Massachusetts Institute of Technology by the 'Special Interest Group in Information Theory' brought together experimental psychology, theoretical linguistics and the computer simulation of cognitive processes (Miller, 2003, p.143). During that symposium, Miller recounts, the usual boundaries between these disciplines were set aside to make space for a novel vision of the human being: one in which invisible mental processes and the internal life of the mind were considered the foundation of our actions

and behavior, and which moreover were assumed to be grounded in a *biological and measurable reality*. This was a radical departure from behaviorism – the dominant theory at the time that rose to popularity since the 1920s – that had hitherto firmly denied the significance of mental processes, unless they were manifested in observable behavior (Watson 1913; Skinner 1953, 1974; Graham 1984; for a review see Johnson and Erneling, 1997). For behaviorists, the 'black box' of the human mind was irrelevant if it was not directly linked to behavior, in the sense that any 'internal' psychological hypothesis had to correspond to 'external' behavioral data in order to be considered valid[11]. In this sense, behaviorism's impact on the psychological sciences of its time was that it defined it as a science of behavior, not as a science of the mind, and that therefore, it was ultimately assumed that human behavior could be described and explained without final references to internal mental states (Graham 2015; Zuriff 1985). Regarding theories of the socio-political world this meant that social behavior had to be explained based on the behavior itself, and not by tracing the actual cause of behavior back to individual perceptions or beliefs. In the words of behaviorism's most famous proponent, B.F. Skinner, "the objection to inner states is not that they do not exist, but that they are not relevant in a functional analysis" (Skinner, 1953, p.35).

The 'cognitive revolution' of the 1950s therefore positioned itself in direct opposition to behaviorism's dismissal of the mind's relevance in understanding social behavior, even though explaining complex social or even political behavior was not yet on the radar of the revolution's pioneers. Rather, as linguist Noam Chomsky, a prominent anti-behaviorist at the time remarked laconically, defining psychology as the science of

[11] For a more nuanced discussion of different schools within behaviorism see Graham (2015).

behavior was like defining physics as the science of meter reading (Chomsky & Rajchman, 2006). There had to be more to the internal processes of our minds than the speculative mystery surrounding it, yet in order to buttress such a claim, the 'cognitive revolution' needed to assume that representational and computational capacities of the human mind existed, and that furthermore, these were structurally and functionally realized in the human brain, and come up with concrete evidence to prove it (Miller, 2003, p.144). In other words, the brain's 'black box' had to be assumed to have an underlying, discernible and measurable structure and pattern that could be cracked; the human mind and its intricate workings had to be placed at the center of the inquiry into human behavior in a social world.

In the U.S., the involvement of the Alfred P. Sloan Foundation in 1976 was a turning point for what came to be known as 'cognitive science', as well as for its arguably most prominent subfield, neuroscience. According to some of the main actors on the cognitive front at the time, at least six disciplines were involved in establishing 'cognitive science': psychology, linguistics, neuroscience, anthropology, computer science, and philosophy. (Keyser, Miller, & Walker, 1978). These disciplines in turn cross-pollinated each other, resulting in eclectic interdisciplinary subfields such as cybernetics, psycholinguistics, computational linguistics, and evolutionary anthropology, amongst many others. Indeed, the Sloan Foundation's main goal was to unify differing disciplines towards the shared goal of bridging the gap between the mind and the brain. Even though the Sloan Foundation initiative did not carry on beyond the 1980s, its impact was lasting and

started an ongoing, multidisciplinary effort into deciphering and explaining mental processes[12].

In the case of cognitive neuroscience, this effort was premised on uncovering the neural basis of perception, affect, language understanding, decision-making and memory (McClelland and Ralph, 2015; Bennett and Hacker, 2013). The discipline had its major breakthrough in the 1990s with the increasing availability of electroencephalography (EEG), facial electromyography (fEMG), positron emission topography (PET) and functional magnetic resonance imaging (fMRI) methods, which for the first time allowed social psychologists to access the neural mechanisms underlying mental processes.

It is important to note that in the U.S. research context, access to these hitherto 'hidden' neural mechanisms coincided with the Civil Rights Movement and its aftermath, in the course of which social norms prohibiting the overt and explicit expression of racial prejudice began to emerge (Harmon-Jones and Winkielman, 2007, p.4). Social psychologists therefore became increasingly interested in using novel methods to detect implicit and unconscious biases, since explicit self-reporting methods were potentially hiding and obscuring people's social prejudices. The initial research, using physiological measures such as sweat gland activity and skin-to-skin contact, focused on Black-White racial relationships, as well as the faster automatic nature through which Black people were perceived and judged negatively (Cooper and Siegel, 1956; Rankin and Campbell, 1955; Vidulich and Krevanik, 1966); more recent research used fMRI methods to observe

[12] See Vauclair and Perret (2003) for an alternative historical account of the 'cognitive revolution' from a European perspective. The authors contest Miller's (2003) exclusively U.S.-based account and point out Europe's contribution in pushing for a developmental perspective of the human mind and brain, based on the work and efforts of psychologists such as Jean Piaget and Lev Vygotsky.

White participants' distinct brain activity when perceiving Blacks (Hart et al. 2000; Phelps et al. 2000; Amodio, Kubota, Harmon-Jones, and Devine, 2006).

In contrast, within Europe, what came closest to the Civil Rights Movement in terms of socio-political saliency was the stupendous success of European fascism and the Holocaust. Social theorists such as Theodor W. Adorno, a Jewish-German émigré who returned to post-War Germany to establish the *Frankfurt Institute for Social Research*, used innovative psychological methods to determine what kind of personality type would facilitate the rise of fascism, eventually coining the concept of the "authoritarian personality" (Adorno, Frenkel-Brunswik, Levinson, and Sanford, 1950). Adorno wanted to understand why a large swath of Germany's population completely submitted and subserviently devoted itself towards a single dictator and his party, and moreover, how the population could turn into helpful accomplices in executing the chilling plan of exterminating its fellow Jewish citizens. However, Adorno's return to Europe after WWII was unusual, and indeed a considerable part of his research was developed during his exiled time in California with U.S. psychologists such as Nevitt Sanford (Altemeyer, 2004, p. 85). Subsequent research expanded the inquiry into the "authoritarian personality" by looking at right-wing authoritarian personalities and the role of domineering personality traits in politics (Altemeyer, 1981; Eysenck and Eysenck, 1976; Christie and Geis, 1970), as well as submissiveness and obedience to authority (Milgram, 1974; Blass, 1999; Haney, Banks, & Zimbardo, 1973).

Overall, the majority of the pioneering research inspired by the genocidal and fascist events in Europe took place largely in the U.S., with some few exceptions such as Henri Tajfel's contributions to experimental social psychology in the U.K. (Tajfel, however,

was a Jewish refugee from Poland). Amongst U.S. social psychologist who studied the interpersonal and cognitive dynamics behind mass atrocities, there were a considerable number of Holocaust second-generation or child survivors, such as Herbert C. Kelman (1989) at Harvard, Ervin Staub (1989) at Stanford, and Peter Suedfeld at Illinois University (before moving to the University of British Columbia), whose research questions were driven by the quest to uncover the mental dispositions and mechanisms behind the socio-political catastrophes they themselves and their families had survived.

The point I wish to make here is this: when we trace the historical path from the 'cognitive revolution' to the birth of social neuroscience, we need to be aware that even though the 'cognitive revolution' provided breakthrough methodologies such as fMRI in accessing internal mental processes and, most importantly, ushered forward a paradigm shift following behaviorism's grip on the social and psychological sciences, it is impossible to truly understand the birth of social neuroscience in the U.S. context without taking into account the *two major historical events* that drove the initial moral and political concerns behind the social brain research, namely the Civil Rights Movement and – more indirectly through the experience of Jewish European refugees – fascism and the Holocaust. In this worth noting here that in this sense, social neuroscience shares a pivotal historical backdrop with 20th century Western political theory and its post-War (Adorno 2005 [1951]; Arendt 1951; Strauss 1953; Habermas 1981, 1992; Rawls, 1971) and post-Civil Rights Movement authors (Appiah, 1996; Koppelman, 1996; Kymlicka, 1995; MacKinnon, 1987; Marshall, 1965; Nussbaum, 2002b; Shelby, 2005; Taylor, 1994; Waldron, 2000; Wu, 2002). This suggests that an interdisciplinary rapprochement between social neuroscience on the one hand, and ongoing debates initiated in 20th

century political theory on identity, diversity and rights on the other, is a potentially fruitful endeavor.

The beginnings of social neuroscience were not aimed at reductionism of the social world but marked by an acute awareness of the significance of watershed events, and what they revealed about interpersonal dynamics, group identity and racial norms. It is important to keep this in mind when we apply social neuroscience results to questions within political theory. Although political theorists are justified to proceed with a healthy dose of skepticism and caution towards the brain sciences, social neuroscience as such – both in its initial beginnings and in its current research questions – is not a purely reductionist or determinist enterprise that tries to diminish the complexity of the social world around us. Social neuroscience emerged not simply because of the availability of new innovative methodology, but as a reaction to and absorption of the moral, normative and psychological questions raised by salient socio-political events.

As much as the 'cognitive revolution' had paved the path for social neuroscience, the stress on 'social' by the latter discipline was also a clear move away from the cognitive revolution's idea of the human brain as a solitary computer. Critics of the 'cognitive revolution' prior to and during the 1990s had pointed out that the locating of all meaning- and decision-making, as well as understanding and language acquisition within the individual was problematic, and that we needed instead to approach these mental processes as social achievements (Erneling, 1997, p.275).

For example, psychologist Jerome Bruner, in his book *Acts of Meaning* (1990), criticized the 'cognitive revolution' for overlooking the role of culture in meaning making, and further pointed to its inability to explain variation in meaning making in

different societies. Social neuroscience therefore tried to overcome the 'social' blind spot within the initial 'cognitive revolution' through an expanded idea of the individual. One of the founders of social neuroscience, John T. Cacioppo, in his preface to the reader *Social Neuroscience: People Thinking About People* (2006), puts it poignantly like this:

> "The dominant metaphor for the scientific study of the human mind during the latter half of the twentieth century was the computer, a solitary device with massive information-processing capacities. At the dawn of the twenty-first century, this metaphor seems dated. (…) Just as computers have capacities and processes that are transduced through but extend far beyond the hardware of a single computer, the human brain has come to be recognized as having evolved to promote social and cultural capacities and processes that are transduced through, but that extend far beyond, a solitary brain."
>
> (Cacioppo, Visser, and Pickett, 2006, p.xi).

However, even though social neuroscience consciously adopts a non-reductionist awareness of the social complexities underlying individual mental processes and decision-making, the discipline is also defined by a clear commitment to a non-dualist worldview – something that might not sit comfortably with political theorists and philosophers of mind who believe in a dualistic notion of the mind/brain (Descartes, 1641/1996; Berkley, 1710/1998; Hume, 1739/2000; Penelhum, 1970; Foster, 1991; Penrose, 1990; Chalmers, 1996; Swinburne, 1997, 2014). In other words, social neuroscience unapologetically contends that all mental processes are part of our physical and biological world, whereas dualists believes that mental processes and biological processes can belong to different realities, or consist of different 'substances' (as in 'substance dualism' as first defined by Descartes). Social neuroscience instead tries to expose the neurobiological underpinnings of social information processing, perception and decision-making, always assuming that this neurobiological reality exists and is to a large extent measurable.

That being said, social neuroscientists would concede that the uncovering of this neurobiological reality often does not offer a complete or absolute explanation as to why and how social phenomena occur. In this light, consider the following:

> "Social neuroscience is an integrative field that examines how nervous (central and peripheral), endocrine, and immune systems are involved in sociocultural processes. Social neuroscience is nondualist in its view of humans, yet it is also nonreductionistic and emphasizes the importance of understanding how the brain and body influence social processes, as well as how social processes influence the brain and body. [It is a] comprehensive attempt to understand mechanisms that underlie social behavior by combining biological and social approaches."
>
> (Harmon-Jones and Winkielman, 2007, p.X)

The concept of the social brain is therefore as much an outgrowth of the 'cognitive revolution' as it is a repudiation of the ultraindividualist and computer-metaphor-driven beginnings of this very revolution. The *Oxford Dictionary of Psychology* (Colman, 2006) mentions a 2001 conference held at the University of California, Los Angeles, as a founding moment for social neuroscience, where this new field moved away from a hitherto health-and-animal focused research agenda that had primarily analyzed how the social world affected the peripheral nervous system and other bodily systems (Cacioppo, 1994; Berntson, Sarter, and Cacioppo, 1989; Insel and Winslow, 1998; Panksepp, 1998). Instead, the new field of "social cognitive neuroscience" (Lieberman, 2000; Ochsner & Schacter, 2000) steered towards a more socially informed research agenda that began to tackle topics such as intergroup dynamics and interactions, social perception and inference, mirroring, empathy, racism, sexism, stereotyping, social rejection and self-processes, amongst others (for reviews see Amodio & Frith, 2006; Blakemore, Winston, & Frith, 2004; Harris, Fiske, & Todorov, 2006; Lieberman, 2010; Ochsner, 2007). If "the brain's default focus is social", then

> "who we are as humans has a lot to do with what happens between our ears. What happens between our ears has a lot to do with the social world we traverse, engage, and react to. The former has been the province of neuroscience and the latter the province of social psychology for nearly a century. Recently, scientists have begun to study the social mind by literally looking between the ears using the tools of neuroscience. Social cognitive neuroscience uses the tools of neuroscience to study the mental mechanisms that create, frame, regulate, and respond to our experience of the social world. On its worst days, social cognitive neuroscience is phrenological, cataloguing countless brain regions involved in the vast array of social processes. On its best days, social cognitive neuroscience enhances our understanding of the social mind as well as any other method."
>
> (Lieberman, 2010, p.143)

On the methodological challenge of whether social neuroscience methods are capable of illuminating a truth about the social mind that might be escaping psychologists and social scientists, John Cacioppo and his colleagues further add to Lieberman's above claims that

> "theory and methods on social neuroscience can draw upon evidence from the neurosciences to constrain and inspire social psychological hypotheses, foster experimental tests of otherwise indistinguishable theoretical explanations, and increase the comprehensiveness and relevance of social psychological theories"
>
> (Cacioppo, Lorig, Nusbaum, & Berntson, 2008, p.399)

What transpires in the above assessments by social neuroscientists of their own field is that implicit, automatic social cognitive processes that were hitherto considered invisible, inexistent and (most importantly) unmeasurable by political psychologists and social scientists can actually be measured through fMRI and EEG techniques, and hence contribute to understanding the nature of the phenomenon itself. Lasana Harris, Susan Fiske and Alexander Todorov (2006) call this the dual-process perspective (Chaiken & Trope, 1999), in which one is aware of both deliberate and explicit social cognition processes as well as implicit and automatic ones – effectively understanding social cognition as an interplay between both deliberate and non-deliberate, implicit responses at the brain and behavioral level. Social neuroscience's conceptualization of the human mind thus offers political scientists a more complete, dual-process-based idea of the

political self in relation to itself and others, in which the multiple strands of an individual's cognitive, affective, nervous, endocrine and immune experiences at both the explicit *and* implicit level are acknowledged in their entirety. Methodologically, social neuroscience offers additional clarification on this relationship, by being able to show how implicit processes sometimes affirm but also often contradict what is being expressed at the explicit level (Chartrand & Bargh, 1996; Chiao et al., 2009).

In addition, in Harris, Fiske and Todorov's opinion (2006, p.76), social neuroscience in the beginning was significantly more influenced by developmental psychology than social psychology, in that the first attempts at studying the neural models of social cognition began within the field of primatology rather than social psychology (Brothers, 1990; Gross et al., 1972; Perrett et al., 1982), and further, that initial research by burgeoning social neuroscientists (Fletcher et al., 1995; Frith & Frith, 1999) first drew from important insights into the development of social cognition features such as 'Theory of Mind' during childhood (Wimmer & Perner, 1983) and from the methodological advantages of brain lesion studies – two areas that social psychology has not traditionally show much conceptual or methodological interest in.

In summary, the 'social' turn of the 'cognitive revolution' offers political scientists and theorists a more organic (i.e. incorporating developmental social cognition during childhood and adolescence) and more complete (i.e. studying both explicit and implicit social cognition) idea of how we function as social beings. This is made possible through new methodological techniques available to us, which come with benefits and drawbacks, both of which shall be examined later in this chapter. The challenge for political scientists here is to judge where and how social neuroscience's more sterile,

laboratory-tested concept of social cognition can help enlighten empirical and normative problems in the messiness of the political world.

2.3. Political Neuroscience: Have We Finally Arrived?

Political science is no stranger to biology and the life sciences. In the 1980s, Thomas Wiegele founded the Center for Biopolitical Research and the Association for Politics and the Life Sciences (APLS) at Northern Illinois University, from where two notable students, Robert H. Blank and Samuel M. Hines Jr., continued the initial work through interdisciplinary publications (Blank and Hines, 2001; Hines, Funke, Losco, Stewart, & Blank, 2014). However, attempts at bridging the two disciplines remained largely confined to a few authors, and research initiatives in this new field were not widespread and significant enough to change methodologies, concepts or paradigms within the political science field over the last decades as a whole.

One could argue that this lack of interest might be a reaction to the sinister outcome coming out of the merging of politics and biology during colonialism and slavery (McCarthy, 2009), as well as to global eugenics and genocidal programs during the 20th century (Benz, 1999; Hansen & King, 2013; Ihrig, 2016; Tatum, 2010). These historic events had made political scientists cautious to adopt interdisciplinary biological approaches that use evolutionary theories, behavioral genetics or neuroscience methods to explain the outcome of political behavior, especially when this entails moral and normative implications. John R. Hibbing and Kevin B. Smith, two political scientists who have been trying to advance a closer relationship between political science and the life sciences, attest that

> "virtually the entire research agenda in political science is isolated from the vast biological knowledge base that has built up over the course of the past fifty years. Most political scientists continue to be environmental determinists, believing that human behavior is entirely the product of environmental forces. From this perspective, political attitudes and political behavior are driven by parental socialization, campaign messages, conversations at work, and idiosyncratic experiences. Political science gives biology virtually no role in answering the questions it seeks to address. In the extant literature of mainstream political science, biological independent variables are extremely rare, and biological theory rarely is used to generate hypotheses and insights."
>
> (Hibbing & Smith, 2007, p.6-7)

Hibbing and Smith believe that one reason for the lack of interest in biology is the particular way in which most current political behavior is couched in "the structure and organization of mass-scale social life", which "seems so uniquely human, so cerebral, and so rational, [that therefore] scholars of politics are prone to conclude that it somehow transcends biology" (ibid., p.8). It is this mass-scale aspect of modern politics that poses a challenge to more simplified explanatory models of social behavior within the biological sciences, including social neuroscience.

It is worth noting here that Hibbing and Smith juxtapose what they consider the "unique humanness" and "rational" character of modern politics against the purportedly animalistic and irrational aspect of human biology – a commonly drawn dichotomy by political scientists. I claim that this is a misleading dichotomy and wrongly caricatures human biological behavior as one-sidedly brutish, unrefined and lacking rationality, whereas in fact, as we shall see in the subsequent chapters, (de)humanizing fellow human beings is an innate social cognitive capacity that we all share, and which has underlying "rational" incentives in the evolutionary context of cooperating and competing in small groups (Pagel, 2012). If political science wants to embrace the 'biological turn', it might have to critically revisit previous assumptions about what constitutes human rationality

and rational political behavior (Kahneman & Tversky, 1979), and how they relate to modern mass politics.

In Hibbing and Smith's opinion, the most persuasive way to convince political scientists to adopt biological concepts and methodologies is to produce "original empirical research". However, as we shall see in the methodological discussion to follow, the solution is not that straightforward. As I argue in Chapter 1, the brain data itself is not self-evident in terms of its implications for empirical political research and theory-building. Although producing original interdisciplinary empirical data is doubtless the most fundamental, indispensable basis for building a new subfield such as political neuroscience, mere production of raw data does not guarantee the interest or acceptance by more mainstream disciplines within social science. In fact, if a brain-based method such as fMRI is presented as the ultimate and superior way to access human cognition and behavior, then as a result, social scientists using different methods might reject the newcomer altogether, not least out of a sense of threat to their own methodological commitments and research identities (Theodoridis & Nelson, 2012, p. 28). In order to overcome "reflexively dismissive" reactions by political scientists towards neuroscience, there needs to be an "interaction of the two levels" (i.e. political psychology and neuroscience) and not claims of superiority of neuroscience above the rest (ibid.). This is echoed by calls to use neuroscientific contributions to "build on rather than substitute for the extant theory and methods in political psychology" (Cacioppo & Visser, 2003, p.655).

Aside from cross-disciplinary defensive reactions amongst researchers themselves, there are other, intellectually serious reservations voiced by political scientists worth considering (Tingley, 2006; Theodoridis & Nelson, 2012), next to more

enthusiastic voices that support the emergence of 'political neuroscience' as a new field (Fowler & Schreiber, 2008; Jost, Nam, Amodio, & Van Bavel, 2014; Lieberman, Schreiber, & Ochsner, 2003; McDermott, 2004; 2009). I shall now turn briefly towards the position of the latter camp, before moving on to the problematic methodological issues surrounding the interdisciplinary merging between social neuroscience and political science.

In a recent review of political neuroscience as a subfield, John T. Jost and his colleagues (2014) at NYU describe the state of the field as the "beginning of a beautiful friendship" – a departure from John Cacioppo and Penny Visser's (2003) less optimistic diagnosis of the relationship between political psychology and social neuroscience as *not* being quite yet "comrades in arms". A decade after this diagnosis, Jost et al. express excitement at being able to use neuroscientific methods to access the neuronal basis for political partisanship, ideological affiliation, decision-making and voting behavior (Knutson, Wood, Spampinato, & Grafman, 2006; Marcus, 2000; Westen, Blagov, Harenski, Kilts, and Hamann, 2006; Amodio, Jost, Master, & Yee, 2007; Zamboni et al. 2009; Roy, 2016). In addition, political neuroscience draws heavily on social neuroscience studies on interpersonal perception, prejudice, racial bias, and exclusion (for reviews see Cunningham & Van Bavel, 2009; Derks, Scheepers, & Ellemers, 2013; Kubota, Banaji, & Phelps, 2012).

Methodologically, Jost et al. believe that neuroscientific methods can offer greater objectivity by avoiding social desirability and self-representational biases that frequently occur with self-report and survey-based methods, both of which are commonly used in political psychology and political science research. Apart from being able to access

politically and interpersonally relevant neuronal mechanisms through fMRI and other neuroscience methods in the context of specific tasks or situations, Jost et al. discuss the wider project of political "brain mapping", which consists of detecting correlations between region-specific brain activation and particular political behaviors or attitudes.

For example, the amygdala and insula are implicated in racial perception (Amodio, Harmon-Jones, & Devine, 2003; Hart et al., 2000; Kaplan et al., 2007; Phelps et al., 2000), with greater amygdala activity in Whites' responses to Black than same-race faces (Krendl, Macrae, Kelley, Fugelsang, & Heatherton, 2006; Ronquillo et al., 2007; Wheeler & Fiske, 2005). Meanwhile, the medial prefrontal cortex (mPFC) and temporo-parietal junction (TPJ) are active during impression-formation of others, mentalizing, "Theory of Mind" and self-referential processing – politically speaking, this can translate into heightened mPFC and TPJ activity when politicians violate voters' expectations because it requires voters to mentalize the politicians' intentions behind the violation (Cloutier et al., 2011). In another study, the TPJ is more active during collectivist (vs. individualist) concerns because of the need to consider the perspective of fellow society members (Zamboni et al., 2009). In a study on political candidate preferences, the 'ventral striatum' brain area emerges as relevant during reward and value processing (Gozzi et al. 2010; Tusche et al., 2013), whereas the anterior cingulate cortex (ACC) is both implicated in partisanship (Kaplan et al., 2007; Westen et al., 2006) and greater sensitivity to conflicting information and tolerance of ambiguity in liberals (Amodio et al., 2007; Kanai et al., 2011).

In the larger context of a political "brain mapping" project, one can imagine that the above studies could tempt researchers to treat specific brain regions as the 'ideology

brain area' or the 'partisan brain region', or even claim that we can draw a distinction between a "conservative brain" vs. a "liberal brain" – which is exactly what Hibbing, Alford and Smith (2014) proposed in their controversial paper on the conservative "negativity bias". Based on the vast evidence on preconscious and implicit biases, as well as the psychological and physiological differences along individuals' political orientation, Hibbing et al. claim that there exists a cognitive "negativity bias" amongst conservatives, which primes them to respond more strongly and with higher sensitivity to negative stimuli and events. As a consequence, conservatives will try to evade new, potentially threatening information (Castelli & Carraro, 2011; Shook & Fazio, 2009) and ambiguity (Young, 2009), focus on negative information (Dodd et al., 2012), endorse public policies that minimize threats, support strong political authority figures (Altemeyer, 1981; 1996), resist change and defend social traditionalism (Jost et al., 2003; Schwartz, Caprara, & Vecchione, 2010).

The clear-cut alignment of political orientation with more deep-seated psychological, neuronal and physiological individual features seems alluring but also potentially highly problematic on many levels. Indeed, in the open peer commentary on Hibbing et al.'s paper, skeptical voices abound: Cacioppo, Cacioppo, & Gollan (2014) point out problems in how the "negativity bias" is conceptualized, namely that it conflates multiple potential causes that are producing the bias; Evan Charney (2014) believes that the authors' definition of 'conservatism' lacks the distinctive political and economic features that usually distinguish conservatives from liberals, thus rendering their claims about the cognitive differences between liberals and conservatives meaningless; similarly, Huddy & Feldman (2014) question Hibbing et al.'s single-

dimensional characterization of political ideology along left-right or liberal-conservative lines, stressing the multidimensional and diverse origins of political ideology; Olivola & Sussman (2014) argue that many behavioral tendencies ascribed to 'conservatives' are in fact malleable and bear no connection to negativity; Janoff-Bulman & Carnes (2014) point out that a presumed liberal person's failure to avoid a negative stimulus is not the same as approaching a positive one; and finally, cultural neuroscientists Pornpattananangkul, Cheon, & Chiao (2014) draw attention to the heterogeneous nature of negativity bias across cultures and its domain-specificity, based on their own fMRI-based research on the relationship between "social dominance orientation" and empathy (Chiao et al., 2009).

What transpires in these responses by leading political psychologists and neuroscientists is a cautious attitude towards political "brain mapping" projects that make ontological claims (e.g. "this is who conservatives are at the brain-level") and attempt to reduce a complex phenomenon such as an individual's ideological commitment to a single bias trait. It is worth noting that a large majority of the 26 open peer comments applaud Hibbing et al. for their interdisciplinary hypothesizing overall, support their attempts at showing the "limitations of the rational view of the political mind", and agree with the need to demystify and destigmatize 'conservative' political attitudes by understanding their underlying biological and cognitive motivations.

Cracking the 'black box' of the political brain therefore has to be embarked on with methodological and conceptual caution, and any "brain mapping" project has to refrain from over-simplification of political phenomena for the sake of less ambiguous "mapping results". Political neuroscientists also need to be transparent about the fact that

no single brain region is responsible for one sole task, affect or attitude, but that we need to focus on the neuronal *networks* (and hence multiple brain regions) implicated in any one task (Faingold & Blumenfeld, 2014). Additionally, just because we know that a certain brain region is usually engaged in a certain function, this does not automatically imply that activation of this brain region must mean the presence of a particular mental process (also called 'reverse inference') (Poldrack, 2008).

With all these caveats in mind, I argue that the new field of political neuroscience still holds immense potential and excitement for political scientists and theorists alike. Political thinkers have been wondering since antiquity about the cognitive abilities and mechanisms underlying political behavior, resting their assumptions on religious and spiritual beliefs, speculative theories about the human nature of the political animal, and more recently, behavioral observation and attitudinal data collection methods. With neuroscientific techniques, we can start to gain an unprecedented glimpse into how the human brain works in specific social and political situations, even though this glimpse is (and might forever remain) incomplete in terms of grasping the global entirety of social brain functions.

The most important issue for political scientists is that we must be clear about what kind of political question we are bringing to the brain data, because the latter is not self-evident in its political meaning and implications for implementation in the political world. In order not to lose epistemic, conceptual and normative sovereignty over interdisciplinary endeavors at the nexus of neuroscience, the life sciences and political science, it is crucial that political scientists and theorists are able to formulate the unique political question that they want to have answered through brain-based techniques such as

fMRI, and be aware of both the capacities and limitations of these techniques for their specific political objectives.

Earlier, Cacioppo and Visser (2003) mentioned the need for a multilevel integrative analysis, which doubtless is a vital cornerstone for successful interdisciplinary empirical and theoretical work done in political neuroscience. Echoing this, political scientists Alexander Theodoridis and Amy Nelson contend that the

> "(…) future role of neuroscience in political psychology need not be viewed as a higher standard of proof looming on the horizon that will transform work done to date into early theoretical building blocks. Rather, we believe neuroscience has the potential to be integrated into political psychology primarily as a window into a different level of analysis"
>
> (Theodoridis & Nelson, 2012, p.28)

In addition to this I argue that we also need to be very clear about the relationship between these different levels of analysis, and exactly how the neuroscientific level can help address the political question we bring to the table. This stands somewhat in contradiction to Hibbing et al.'s (2007) belief that what political neuroscience needs in order to be accepted by the discipline of political science as whole is simply the production of more experimental data. I argue instead that until we have not figured out how the brain data can relate in epistemic, conceptual and normative terms to the distinct political aspect of the question we are asking, the mere production of more brain-based data does not guarantee acceptance by political scientists and theorists. Developing the PBP for this is a first step towards clarifying the relationship between the two fields. Before we move on to further methodological discussion, below is a summary of the brain-based techniques in question.

2.4. Methodology

2.4.1. Social Neuroscience Methods: EEG, fMRI, and Lesion Studies

The story of Phineas Gage is possibly the most prominent case in the history of neuroscience, and was decisive towards the development of present-day neuroscience techniques such as lesion studies and fMRI. Gage, a young U.S. railroad construction worker from New Hampshire, had a tamping rod violently blown through his eye and skull during a dynamite blast accident in 1848, which in turn severely damaged the orbitofrontal and ventromedial cortex sections his brain[13]. His doctors did not expect him to survive, even less to ever regain consciousness again. However, in the course of a year, Gage seemed to initially be able to resume his former life, returning to work and engaging in relationships with his friends and family. The brain damage that he suffered in the above specified regions however changed his behavior and personality; friends reported that the formerly friendly and reliable Gage had become irresponsible, antisocial and arrogant, even though from a motor-sensory viewpoint he was functioning as before the accident (Cacioppo et al., 2003; Damasio et al., 1994; Harlow, 1868; Macmillan, 1992). Gage's brain injury was the obvious explanation for this shift in personality, which subsequently allowed scientists to make causal inferences about the involvement of certain damaged brain regions' in executive decision-making, impulse control and perspective taking.

Gage's case was the first compelling example in the recorded history of neuroscience where a specific brain region could be ascribed to cause certain behavior. Although this inference might seem obvious to us today, Gage's case was a powerful proof against fluid-based theories of brain function and the nervous system that circulated

[13] However, because no immediate autopsy was carried out after Phineas Gage's death, the exact damaged brain regions are somewhat disputed. See Macmillan, 1986.

up until the mid-19th century (Finkelstein, 2014; O'Shea, 2005), as well as against phrenological theories espoused by Franz Joseph Gall (Macmillan, 1992).

Methodologically, Phineas Gage paved the way for the brain lesion studies of today, which examine living individuals with brain damage in particular areas, and look at the ensuing psychological and behavioral effects. The most important advantage of lesion studies over fMRI is that researchers are able to draw solid causal inferences between specific brain regions and psychological states (Lieberman, 2010; Mazziotta, Toga, & Frackowiak, 2000).

Another 19th century brain injury case that is even more directly related to brain imaging techniques such as fMRI, but less famous than Phineas Gage, is that of the Italian peasant Bertino (Raichle & Snyder, 2000). Bertino had suffered a head injury where parts of his frontal lobes became exposed. His doctor, Angelo Mosso (1881), observed a curious phenomenon: every time the church bells rang, Mosso saw an increase in the magnitude of pulsation over the frontal lobe, which was unrelated to Bertino's heart rate or general blood flow. Mosso probed Bertino further by asking him questions about his emotional state and to solve mathematical problems, all of which elicited increased brain pulsations. The case of Bertino set the stage for hemodynamic (i.e. blood-flow based) measurements of the brain such as fMRI (Cacioppo et al., 2003; Raichle & Snyder, 2000).

Other discoveries, such as for example by Luigi Galvani and Alessandro Volta in the 18th century on the role of electricity in operating the nervous system – which they studied through experimenting with frog muscles (Piccolino, 2013) – later enabled the development of EEG techniques, which measures electrical activity from neurons firing

in the outer cortex. Both fMRI and EEG are two of the most commonly used brain-based measures in social neuroscience today (in addition to less frequently conducted lesion study experiments).

It is beyond the purpose of this chapter to give a complete account of the history of the development of neuroscience methods and techniques, but the selective cases above illustrate the often contingent and accidental nature through which breakthroughs in neuroscientific discovery and methodological advancements were made. Most important, methodologies were developed out of and in accordance with the limitations and restrictions of actual human subjects. Today, even though neuroscience methods and insights have advanced immensely since Phineas Gage and Bertino, researchers still face the 'human subject limitation' challenges, i.e. how to justify generalizations made based on one's sample subjects, how to ensure that subjects adhere to the experiment script, how to prevent subjects from preempting the researchers' study motives, how to ensure that subjects receive a shared stimuli in the same fashion, how to pursue research questions without violating ethics codes and subjects' privacy, etc. (Lieberman, 2010). Political scientists who are skeptical of neuroscience or psychology-based methods need to be aware that neuroscientists face a wide array of methodological constraints and humbling challenges before and during experiments, as well as afterwards, during data analysis and when drawing theoretical conclusions.

So finally, what is 'functional magnetic resonance imaging', also known as fMRI? The simple answer is that fMRI measures differences in blood flow to brain regions during specific tasks or exposure to stimuli. Blood oxygen level-dependent (BOLD) fMRI is based on the idea that blood flowing to an active region is more

oxygenated than blood in inactive regions, and that further, oxygenated blood has different magnetic properties than deoxygenated blood in terms of its hydrogen molecules (Lieberman, 2010; for reviews see Faro & Mohamed, 2006; Heeger & Ress, 2002; Jezzard, Matthews, and Smith, 2001). FMRI takes advantage of the intrinsic magnetic moment of these hydrogen molecules (which behave like small magnets when placed in a magnetic field) by introducing a second magnetic field that oscillates at a particular frequency, which in turn causes the hydrogen molecules in the examined brain to rotate around the direction of the larger, stable field. It is this rotation that creates a detectable signal, which the fMRI machine picks up spatially and reconstructs it into three-dimensional brain imaging pictures.

Unlike MRI, which is primarily used in medical diagnosis in depicting the anatomical features of the brain, *functional* MRI offers brain imaging during dynamic processes, i.e. whilst the brain is completing certain tasks or reacting to a string of stimuli. The fact that fMRI is a non-invasive technique that can be conducted with healthy individuals is perceived as somewhat of a methodological revolution for researchers working in the fields of cognitive and psychological sciences:

> "In decades past, studies of the neurophysiological structures and functions associated with psychological events were limited primarily to animal models, postmortem examinations, and observations of the occasional unfortunate individual who suffered trauma to or disorders of the brain. Developments in electrophysiological recording, brain imaging, and neurochemical techniques within the neurosciences have increasingly made it possible to investigate the role of neural structures and processes in normal and disordered thought in humans. The importance of these technical developments was underscored by Congress's declaration of the 1990s as the decade of the brain."
>
> (Cacioppo & Berntson, 1992, p. 1020)

The more complicated answer would be that fMRI measurements are in fact not a direct reflection of neuronal activity (even though we commonly call the results "brain

activity") since fMRI measures hemodynamic rather than neural responses. However, as others have pointed out, hemodynamic responses are strongly reflective of synaptic (i.e. neuronal) activity (Roland, 1993), which has been shown through parallel test with additional methods such as EEG, local field potentials, multiple unity activity, and laser Doppler flow measurement (Cacioppo et al., 2008; Heeger & Rees, 2002). The complication therefore lays not so much with the indirectness of neuronal activity measurement but in the fMRI procedure and statistical analysis itself.

For example, although fMRI is excellent in offering a fairly accurate *spatial* picture of where blood flow activity is located, it lags behind on *temporal* resolution, in the sense that hemodynamic response takes about 2-4 seconds to reach its peak in response to a stimuli or task, and then another 6-12 seconds to decline, therefore being unable to track brain activity on a more detailed, millisecond by millisecond basis (Cacioppo et al., 2008). Although fMRI is already offering a great temporal improvement to older social neuroscience methods such as Positron Emission Topography (PET) scans, which comparatively had far slower temporal resolutions (roughly one aggregate data point per minute), fMRI cannot rival EEG in its millisecond-based temporal results of brain activity as measured through synaptic electric signals. We shall discuss EEG later in this chapter. In addition, the BOLD signal is dependent on the establishment of a "baseline" rate of oxygen usage from which to compare changes in blood flow. Others have pointed out that this "baseline" or observed "inactivity" in other brain regions does not necessarily mean that nothing interesting is happening in these "inactive" brain regions (Fox et al., 2005; Martuzzi et al., 2010). In other words, both brain "activity" and

"inactivity" need to be interpreted as significant, even though their significance varies in regard to the studied psychological reaction in the experiment.

There are further complications in terms of experimental procedures for fMRI. It is important to note that fMRI data is "preprocessed" before the actual analysis, in that the raw obtained data has to undergo various procedures before it appears as the colored brain imaging researchers print in their research articles (Lieberman, 2010, p.146). fMRI studies are commonly conducted with a very small-n number (on average about 10 subjects). During the scanning process, researchers first have to carry out "realignment" of small movements made by subjects', i.e. recorded brain images have to be corrected for any movements so that the same brain region shows up in the same place during the whole data collection process. Secondly, researchers have to "normalize" the brain scans of all subjects – who naturally have differing brain sizes – into a single coordinate space. Finally, spatial "smoothing" has to be applied, in which "voxels" (i.e. three-dimensional pixels) from the raw brain scans are averaged through statistical analysis. This is chiefly done to strengthen the detection of certain signals that researchers are looking for.

In this complicated "preprocessing" procedure, one can imagine how many things might go wrong. If "realignment" is not carried out correctly, the actual location of crucial brain regions might be indicated in error. In the case of spatial "smoothing", there have recently been concerns voiced over the procedure's statistical reliability and its possible overconfidence in detected signal strength (Vul, Harris, Winkielman, & Paschler, 2009; Eklund, Nichols, & Knutsson, 2016). In other words, neuroscientists who use fMRI methods are criticized for presenting to the research community an overprocessed, somewhat stylized and oftentimes statistically enhanced picture of "brain

activity". Although these concerns are valid, neuroscientists take them seriously (Cacioppo et al., 2003; Lieberman, 2010) and efforts are made in presenting more raw data within research reports and point to statistical weaknesses.

The latest, most stinging critique (Eklund, Nichols, & Knutsson, 2016) about faulty statistical software used in fMRI "smoothing" analysis had to retract its initial estimate of 40,000 affected fMRI research studies and lowerd it down to about 3,500 studies. The issue there was a bug detected in a specific software program used by fMRI researchers; but more generally, the concern was about the need in the neuroscience community to report complete results rather than selective snippets of "preprocessed" brain images and to share data more widely with colleagues[14].

Although these criticisms are concerning, other fields in the natural sciences, or indeed political science itself, suffer from similar problems in terms of raw data transparency and statistical dependability. In the most prominent recent case within political science, a publication by political scientists Michael J. LaCour and Donald P. Green (2014; retracted 2015) in the eminent natural science journal *Science* on the effect of personal contact with gay canvassers on voters' support for same-sex marriage had to be retracted because of fabricated and manipulated data (Broockman, Kalla, & Aronow, 2015). Although the lead-author LaCour seems to have caused the actual data manipulation and fabrication, the co-author Green (in a more senior research position than LaCour) failed to notice this mistake because he never reviewed the raw data himself, despite agreeing to co-authorship of the paper (Bohannon, 2015 in Science

[14] See discussion of the Eklund, Nichols, & Knutsson (2016) article by one of its authors, Thomas Nichols, on his Warwick University Blog at http://blogs.warwick.ac.uk/nichols/entry/bibliometrics_of_cluster/

Mag)[15]. Therefore, if it is common practice amongst political scientists *not* to necessarily seek access to all the raw data available – even if one is a co-author of the very study that uses the raw data to make claims – then neuroscience's problem with raw data transparency seems to be more common and widespread.

It is worth noting that in both the neuroscience and the political science cases discussed just now, it was fellow researchers who discovered and pointed out these methodological problems. As former *Science* editor-in-chief and current president of the National Academy of Sciences Maria McNutt stated in response to the La Cour and Green retraction, "fortunately, science is a self-correcting process: researchers publish work in the scholarly literature so that it can be further scrutinized, replicated, confirmed, rebutted or corrected. This is the way science advances." (in Bialik, 2015). This is why this chapter argues that fMRI should not be flatly dismissed based on reservations about data transparency and statistical reliability. As important as these reservations are – and will continue to be brought to light by fellow researchers in the neuroscience community and beyond – they are not a powerful enough reason to dismiss the significance and insights gained by fMRI altogether. Statistical bugs can be fixed and will continue to appear, data fabrication will continue to happen as long as there are bad (i.e. dishonest and unethical) apples amongst researchers (which is also exacerbated by the "pressure to publish" amongst all fields in research academia), and greater data transparency is a goal that the scientific research community is working on.

An important point here is that as long as political scientists who wish to appropriate methods from the neurosciences (or draw conclusions from brain data) are

[15] For an interview on research ethics with Donald P. Green, co-author of the study, see http://nymag.com/scienceofus/2015/05/co-author-of-the-faked-study-speaks-out.html

aware about methodological problems and pitfalls, it is easier for them to avoid drawing unrealistic or unsustainable conclusions. The self-correcting process of science should be seen as its inherent strength and a fairly reliable bulwark against methodological mistakes made in brain imaging, at least in the long-term. Most crucially, outsiders to brain imaging techniques have to analyze and draw conclusions from its data based on the premise that fMRI, for example, does not offer causal inferences but only correlational ones, and that just because one brain region "lights up" during a specific task, this does not automatically mean that activity in other regions is completely irrelevant. Apart from these caveats, more practical challenges during fMRI data collection consist in fMRI machine conditions that can induce claustrophobia in participants, selectivity bias due to the fact that experimental participants are predominantly U.S. undergraduates (bias both in terms of age and educational status), and the expensive price label attached to owning and running fMRI machines.

FMRI is not the only technique to measure brain activity. More established and economical techniques are *electroencephalography* (EEG) and *event-related brain potentials* (ERPs). EEG is a "recording of minute electrical changes that occur on the scalp" and ERPs are the same signals but "collected in a different paradigm" (Cacioppo et al., 2008); ERPs are frequently used in a complimentary way together with fMRI studies but also very much independently by itself. The most important difference between EEG and ERPs is that for ERPs, a sequence of stimuli is presented to the subject and the exact moment when the stimulus is presented can be retraced in the data. During ERP recording, subjects wear the characteristic "cap" that is able to record voltage

changes in synapses in the outer cortex – allowing more movement flexibility than fMRI procedures and a less complicated experimental set-up.

The great advantage of ERPs is their excellent temporal resolution – being able to record signal changes by milliseconds – which is crucial for when a study tries to find out exactly when (and how) subjects react to exposure to a political candidate or an other-race face, especially when survey-based responses do not suffice in answering these questions (i.e. did subjects have a certain brain reaction before or during their survey responses?). Moreover, EEG was one of the first methods available to social neuroscientists with which they could show the existence of vital dissociations between social cognitive processes (Cacioppo, Crites, & Gardener 1996) and conduct experimental tests on "otherwise indistinguishable theoretical explanations" (Cacioppo et al., 2008).

Particularly in the field of cultural neuroscience (Ames & Fiske, 2010; Chiao & Ambady, 2007; Gutchess & Goh, 2013; Han, 2015), where subjects can be bicultural (e.g. identifying with both their immigrant and host country identity), the behavioral response or self-judgment can often differ from what is experienced at the brain level. For example, an Asian-American participant might express behaviorally or in their self-judgment a preference for being American or Western (e.g. in order to conform to social expectations of the host country) in the sense of preferring more individualist styles of self-construal, but when studied at the neural level, their cultural self-construal might instead be more Asian, in the sense of being more collectivist (Chiao et al., 2009). When studying individuals who are bicultural, traditional psychology and social science methods can face various obstacles: subjects might not have conscious access to their

cultural self-construal style and therefore are unable to answer related questions on surveys (Kitayama, 2002; Nisbett & Wilson, 1977); they might be hesitant or embarrassed to disclose fully their bicultural identities; or the same survey question could be interpreted differently across cultures (Heine et al., 2002). Most crucially, we canot take for granted that introspection or self-report are always accurate or able to provide a complete picture of the multilayered cultural self. Thus in these instances, fMRI and EEG methods are able to draw on a crucial dimension of bicultural reality that other methods are unable to access.

Finally, another method available to social neuroscience is lesion studies. The first lesion study in the history of neuroscience was, as discussed earlier, the case of the U.S. railroad construction worker Phineas Gage. Lesion studies examine individuals who demonstrate damage to certain brain regions and investigate how the damage affects psychological processes and behavior (Lieberman, 2010). The most decisive advantage of lesion studies over fMRI and EEG is that we are able to make causal inferences, i.e. observed psychological deficiencies or a particular behavior can be attributed directly to the damaged brain region in question. Liane Young and her colleagues (2010) used lesion study methods to examine how damage to the ventromedial prefrontal cortex impaired subjects' ability to take others people's perspective (i.e. 'mentalizing' – more on this in Chapter 3) and therefore led them to make harmful and morally questionable decisions towards others. The primary downside of lesion study methods is the availability of suitable subjects and that brain damage is usually rarely limited to one region only, making it harder to draw confident inferences about specific regions.

2.4.2. Methodological Concerns

One of the most common methodological mistakes made is drawing wrong conclusions about brain function during "reverse inference" (Cacioppo & Visser, 2003; D'Esposito, Ballard, Aguirre, & Zarahn, 1998; Poldrack, 2006). This occurs when a correlation between a particular psychological state and a specific pattern of brain activation is established, and when later, in a different experiment, researchers infer from the same pattern of brain activation that this particular psychological v must be present again – even though it might not be. In other words, we cannot automatically infer 'reversely' about the presence of certain mental states just because the same brain regions are activated. Although the field of social neurosciences as a whole is striving towards being able to make more confident reverse inferences through ongoing experiments and improvement of methods and techniques, it is important to tread with caution in this regard.

Theodoridis and Nelson (2012) warn about this specifically when attempts are made at mapping the "political brain" (Jost et al., 2014; Knutson et al., 2006). Political processes, they point out, are more complex than social ones in that they are more specific and peculiar in their neuronal manifestations – it is "easier to imagine the development of distinctly social, as opposed to political, neural processes and brain regions" (Theodoridis & Nelson, 2012, p.37). In addition, effects that may seem large in the lab with a potentially biased subject population (i.e. easily-influenced college undergrads) may be small if tested in the general population and the complex world of politics. In his review of lab-based psychological experiments, political philosopher Maurizio Meloni (2012) draws attention to the fact that positive interpersonal feelings,

such as empathy, detected in lab experiments might only be short-lived and are unable to withstand the challenges of committing to long-term help towards refugees in the political world, for example.

Another danger in the brain mapping context is the "category error", which consists in searching for brain centers "for guilt, loyalty, and negative moods, [instead of asking] what simple features of these complex psychological functions are being processed by specific networks of neural systems" (Cacioppo et al., 2008). In other words, the "category error" is based on the assumption that there exist singular brain centers for mental and behavioral processes, instead of unique brain networks involving various brain regions (Cacioppo et al. 2002; Sarter, Berntson, Cacioppo 1996). This is connected to "reverse inference", in the sense that in order to avoid the "category error", we need to "distinguish between a nonhomogenous brain in which different regions can influence different mental or behavioral processes, on the one hand, and the hypothesized role of these regions as unique locations of the mechanisms underlying these processes, on the other" (Cacioppo et al., 2003, p.653). But is the idea of the nonhomogenous brain (Uttal, 2001) not a contradiction in itself? Only perhaps if we understand the human brain as a computer – as did some of the initial pioneers of the 'cognitive revolution' – instead of the picture that has emerged in the last two decades: the human social brain is a complex entity that is as much biological as it is social, highly plastic and able to reconfigure considerably within a lifetime, and serving a variety of heterogeneous purposes in the entangled webs of human social interactions and transactions.

The best way to avoid the pitfalls of "reverse inference" and the "category error", as well as other methodological mistakes, is to adopt a "multiple integrative level of

analysis", as briefly mentioned at the beginning of this chapter (Cacioppo & Visser, 2003). The point behind this is the belief that even though all human behavior – including political behavior – could be understood biologically, this does not mean that "biological reductionism yields a simple, singular or satisfactory explanation for complex behaviors, or that molecular representation provides the only or best level of analysis for understanding human behavior" (Cacioppo & Visser, 2003, 649f.). The point of 'multilevel integrative analysis' is to acknowledge the complexity of social and political phenomena as a starting point, and from there approach these phenomena "from various scales or perspectives, such as the neuroscientific, cognitive, social, and political (…) [and further use] observations at one level of organization (…) to inform, refine or constrain inferences based on observations at another level" (ibid.).

This echoes the concerns of Theodoridis and Nelson (2012), who warn of the dangers of rooting oneself in one methodological paradigm and treat fMRI as an exclusive 'higher standard of proof', suggesting instead to use fMRI analyses as a "window into a different level of analysis" (p.28). Jost et al. (2014) call for a "collaborative cross-examination" in order to avoid problems that might arise during "reverse inference", which means that researchers should use data and concepts available at the behavioral level to verify the validity of the neural data, resulting in "a decidedly anti-reductionistic, collaborative approach to science in which psychological and physiological methods and interpretations are regarded as equally indispensable" (ibid., p.28). Meanwhile, political scientist Dustin Tingley (2006) argues that interdisciplinary neuropolitical research would benefit from reciprocal influence, in that shortcomings in neuroscience models on affective and cognitive reasoning, for example, could be partly

overcome through borrowing conceptual distinctions from political science from areas such as voting behavior and candidate preference theories.

Last but not least, in the growing field of *neurofeminism*, neuroscientists and gender study researchers have pointed out the gender-biased processes of knowledge production in the neurosciences, which includes the setup of categories, experimental design, result presentation and analyses of result validity (Fausto-Sterling, 1992; Schmitz & Höppner, 2014). Neurofeminists have suggested a review of methodological procedures that display gender bias, and try to draw attention to "neurosexism" (Fine, 2010) and unconscious gender bias in the research design process and execution. This call for a greater sensitivity to identity diversity is also made in the aforementioned new subfield of "cultural neuroscience", where cultural neuroscientists question the implicit assumptions made by neuroscience researchers in the West, on how Western cultural value systems apply to subjects worldwide (or across the West's own turbulent cultural history, see Backman, 2012), and their omission of the complexity and interaction between different co-existing identities within individuals who identify as bi-or multicultural.

I would like to dedicate the remaining part of this section to a methodological conundrum that I believe is possibly the most profound and perplexing of all the various challenges already mentioned, namely that of the "bottom-up" question (Glimcher & Rustichini, 2004; Hughes & Churchland, 1995; Kihlstrom, 2006; Mitchell, 2010; Wilson, 1998). How is brain activity in the form of blood flow and electric signals distinguishable from behavior? Does the neural level of analysis make up the "bottom" of more comprehensive theories about human social action, and do other levels of analysis such as

attitudinal and behavioral ones build on top of the neural level? Or is this kind of conceptualizing too reductionist and simplistic in regard to the interactive, possibly non-hierarchical nature between these levels?

Critics of the "bottom-up" theory argue for example that by grounding all mental processes at the neural level, we end up denying the existence of psychological processes that might not be representable at the brain level (Kihlstrom, 2010). In other words, these critics maintain that psychological concepts such as desires, beliefs and feelings are separate entities in comparison to neurobiological concepts of these phenomena. This is in line with the earlier mentioned idea of *mind-world dualism*, which believes that certain (or all) mental processes are grounded in a separate reality from the physical one. For political neuroscience, this means that critics of the "bottom-up" approach believe that the political brain cannot be understood without the political mind (Theodoridis & Nelson, 2012, p.35).

This dissertation defends the "bottom-up" approach, and develops the PBP based on the premise that neural manifestations of mental processes surrounding dehumanization, exclusion and mentalizing matter fundamentally for political arguments and methodological frameworks. However, defending the "bottom-up" approach does not automatically mean that psychological concepts about the political mind are useless for making sense of the political brain – in fact, observations made both at the level of behavior and the level of the mind are crucial in determining to which extent certain brain processes are politically relevant. The PBP developed in this dissertation contends that it is possible to insist on the relevance of the "bottom-up" idea, without succumbing to a blanket reductionism of political experience and behavior. To echo a point made by

Cacioppo and Visser (2003) – although it might be possible to eventually reduce all political behavior to processes at the brain level, this does not necessarily mean that brain-level based explanations are the sole or superior level of analysis for understanding all political phenomena. What they do offer us is a crucial and fundamental dimension of social reality that is too important to ignore in any comprehensive attempt at political theorizing about the cognitive conditions for living in hyperdiverse societies.

Another way of highlighting problems around the neural "bottom-up" theory of political behavior is to ask ourselves what exactly is the relationship between implicit cognitive processes and explicit behavior – or to be more specific, between implicit cognitive biases and explicit discriminatory behavior. How does bias, prejudice and exclusion at the brain level lead to behavioral manifestations thereof, if indeed at all? A recent study by Patrick Forscher and his colleagues (2016) investigated through a meta-analysis of over 400 existing studies how implicit bias could be effectively overcome. Firstly, they found that procedures that tackled bias by targeting people's motivations and engaged their mental resources were by far more effective than procedures that were based on threat or made appeals to people's emotions and morals. Most importantly however, they concluded that changes in implicit bias hardly led to changes in explicit bias or behavior.

It is worth noting that one of the research team members on this study was Brian N. Nosek, who together with Anthony Greenwald and Mahzarin Banaji invented the original *Implicit Association Test* (IAT) and coined the concept of implicit bias. Is one of the foremost experts on implicit bias suggesting that implicit bias might not matter politically after all, if in fact it often fails to affect actual behavior and explicit attitudes?

In a subsequent interview with *The Chronicle of Higher Education*, Nosek states that although the connection between implicit bias and discriminatory behavior seems "very weak overall", the IAT might still be able to predict political candidate preferences and other relevant political behavior (Bartlett, 2017).

Nosek does not discuss the possibility that the failure to detect significant correlations between implicit biases and behavior might be due to problems related to experimental set-up and limitations of certain psychological concepts of "political behavior". For example, the restricted timeframe of an experiment might make it difficult to gauge significant changes in discriminatory behavior simply because it is too short, not allowing for a measurement of the effects of incremental bias over time. Also, subjects might be succumbing to "acquiescence bias" in avoiding to admit to and display explicitly biased behavior (Tourangeau & Yan, 2007). Other effects outside of the experimental set-up, such as specific events or stimuli that might reinforce an inherent implicit bias (e.g. a negative news story about or direct encounter with an out-group) and could possibly affect behavior also need to be taken into account.

In addition, the way how we define politically relevant behavior matters as well: if researchers understand "effects on behavior" primarily as changes in very explicit attitudes and blatant discriminatory behavior, they might miss the effects of implicit bias on more subtle excluding behavior such as microaggressions (Sue, 2009; 2010) and subtle dehumanization of out-groups (Leyens et al., 2001). In the field of microaggressions in particular, there exists an ongoing, heated debated about the challenges of measuring microaggressions accurately and convincingly at the behavioral level, since the opaque, ambiguous and subjective nature of microaggressive statements

and actions can be difficult to capture through a standardized methodological system (Sue, 2017; Lilienfeld, 2017). A parallel debate within *behaviorism* shows the complexities of categorizing behavior: some behaviorists defend the molecular paradigm of behavior, which views behavior as discrete and separate units that are linked together to make up more complex performances (Keller & Schoenfeld, 1950; Skinner, 1953); other behaviorists believe instead in the molar paradigm of behavior, which, based on previous theories by John Dewey, Karl Lashley and Gestalt psychologists, postulates that all behavior is continuous and organically linked together (Baum, 2002). Social psychologists who wish to test correlations and causations between implicit mental processes and political behavior need to be aware that political behavior is highly complex in its explicit *and* subtle manifestations, and that the complexity of political behavior sets it apart from more simplistic economic models of reward behavior and decision-making. This is why a sophisticated conceptualization of what constitutes relevant "political behavior" is crucial in experiments, in order to be able to draw politically relevant conclusions from them.

This dissertation contends that exclusionary mental processes that are captured more implicitly at the brain level *are* highly relevant for political behavior, even though various experiments reviewed by Forscher et al. might not invite this kind of conclusion at first glance. This is another example where the psychological and cognitive data is not politically self-evident, and where the PBP can help to illuminate the importance of implicit brain processes for politics. Based on the PBP's "bottom-up" premise, as well as additional insights from neuroscience on how repeated mental procedures (in this case, daily exercise and reinforcement of implicit bias) can significantly change synaptic

functions and even whole brain structures and gray matter (Maguire et al., 2000), this dissertation argues that taking into account implicit exclusionary biases against specific social out-groups is vital for any socially-oriented political theory, especially one that theorizes about how cooperation and solidarity in a hyperdiverse political body politic can be attained.

2.5. Theorizing with a Purpose: Mechanism over Ontology

In Theodoridis and Nelson's (2012) assessment, political neuroscience currently offers "interesting empirical results but [is] light on theoretical implications" for political science (p.28). So far, this chapter tried to show that the theoretical implications for political science methodology are overall significant and can even be profound for redefining what constitutes political behavior, attitudes and implicit mental processes – and the way in which they are intertwined. This being said, this chapter also tried to address in detail how neuroscience results and the methods through which they are obtained cannot be employed without reservations and caveats, and that many methodological problems still pertain to the field as a whole. Theodoridis and Nelson believe that the main reason why political science should embrace neuroscience is the duty to collect as much empirical data as possible (and hope that it might be relevant at a future stage), even if the empirical data might not be useful for political science at this point. Contrary to them, this dissertation contends that *the available empirical data on the social and political brain is sufficient to draw meaningful implications* for political theories about exclusion, cooperation and identity. This why in this final section, I wish to summarize why political science in general (and political theory in particular) can benefit from neuroscience methods and the insights gained from it:

1. Political neuroscience data shows that classic political science methods such as survey-based, attitudinal and behavioral measurements can sometimes miss another dimension of mental processes detectable only at the neuronal level. Neuroscience methods can therefore be useful to political science in gaining a more complete picture of the motivations and intentions behind political actions and decision-making. It is this multilevel integrative analysis (Visser & Cacioppo, 2003) that both political neuroscience and political science can help each other strive towards, and benefit from mutually.

2. Any political theory that makes abstract and universalist assumptions about the political animal and the inner workings of her political mind needs to verify these theoretical assumptions, or at least be able to defend them against alternative empirical evidence. Given that political theorists throughout history have proposed theories about the cognitive conditions underlying political action and interests, and that some theorists even attempted to include the scientific evidence available to them at the time, it is only natural that political theorists today should engage seriously with the evidence emerging from political neuroscience.

3. In this context, this dissertation argues that we are not looking for universalist ontology but ***universalist brain mechanisms***: we are not using neuroscience to build theories about who we *are* as political beings (i.e. ontology) but to understand what our brains *do*. This means that we are trying to identify fundamental mechanisms underlying social cognition processes that can be applied to understanding relevant political phenomena.

4. As a counterpoint to the universalist argument, political theorists should pay attention to emerging subfields within social neuroscience such as cultural neuroscience, which in an unprecedented way allow us to access subconscious and implicit identity dynamics in bi-and multicultural individuals, thus offering a crucial insight into the multilayered cultural influences on political perception and decision-making in a hyperdiverse world.

5. Leaning on this, political science as a whole is struggling with increasing its diversity, and up to date still faces considerable underrepresentation of women and severe underrepresentation of racial and ethnic minorities (Mershon & Walsh, 2016). If this lack of diversity persists, and yet (mostly Western White, and often male) political scientists continue to theorize about the needs and demands of minorities, the experience of racism, and the fate of multiculturalism and cosmopolitanism, then at the very least political scientists ought to inform themselves about the mental processes of bi-and multicultural individuals – which includes the neural level – in order to make convincing and justifiable political demands on how hyperdiverse societies should function.

6. Moreover, insights from cultural neuroscience show that cultural self-knowledge is not necessarily confirmed at the brain level, meaning that during self-report or introspection, minorities themselves might be missing an important dimension of

how their cultural identities affect their perception, attitudes and behavior. Neuroscience methods can thus help access this dimension and contribute to a multilevel analysis of how identity influences politics.

7. We need neuroscience methods in order to understand the underlying neural mechanisms that make up hyperdiverse societies (4. and 5.) and to establish a set of generalizable, basic mechanisms that drive political beings universally (2. and 3.). In addition to constructive objectives, we also need to determine which ***cognitive processes are most destructive and debilitating*** to the politics of hyperdiverse societies. Neuroscience methods can help us detect and specify ***implicit, automatic and rapid cognitive processes of exclusion, prejudice and dehumanization***. Understanding what the most destructive cognitive processes are is essential for constructing a minimal neuropolitical theory of cooperation.

8. The new picture of the political brain challenges dominant political science concepts of the ***rational individual***. Whereas political science theories about human motivations and behavior are still largely defined by hyperindividualism and hyperrationalism (Blank & Hines, 2001), insights into the political brain in terms of its exceptional sociability and sensitivity to social cues, as well as its non-standard "rationality" compel us to reconsider previous paradigms. This dissertation treats the political brain's "flexible social cognition" (Harris, 2017) as an opportunity to rethink previous ideas of human rationality, with a particular focus on dehumanized perception (7.)

Being able to make a distinction between ontological vs. mechanism-based objectives (3.) is crucial for gaining clarity on what kind of neuropolitical paradigm we are committing to. This distinction also plays an important role in epistemological debates in political science. Patrick Thaddeus Jackson, a scholar in international relations (IR), outlines how political scientists in IR divide along the lines of "ontology first" (Klotz and Lynch, 2007; Patömaki and Wight, 2000; Wendt, 1987; Wight, 2006) and those who oppose ontology (Chernoff, 2009; Jackson, 2011), describing the dilemma of choosing between the two as follows: "we appear to have a choice between starting with the world and conforming our methodology to that world, or starting with methodology and thus losing the world as we try to articulate universal standards for scientific research" (Jackson, 2011, p.27). Jackson describes ontology in this context as our "hook-

up" to the world, i.e. the conceptual and philosophical basis on which our claims about the political world are made (p.28). An ontological approach to international politics therefore consists in presuming certain pre-existing characteristics and conditions about the political world, to which one's chosen methodology has to be adapted; it is a commitment to a particular way of believing what the world consist in. For example, Alexander Wendt, another IR scholar and proponent of the constructivist school, famously argued in his paper on the agent-structure problem that there exists empirically unobservable structures in the social and natural world that are nonetheless as real and valid (Wendt, 1987).

Jackson warns that the adaption of methodology to ontological concepts can make us "stuck with techniques and standards designed to respond to the specificity of the object under investigation" (p.27). He believes instead that knowledge production is not separate or subordinate to the world, but fundamentally linked to how the world is. Jackson singles out mind-world dualism as part of the driving force behind ontological commitments – and as a major problem for political science epistemology. He rejects mind-world dualism in favor of a more pluralistic account of political phenomena, where "there are a *variety* of claims about our hook-up to the world, and thus a variety of philosophical ontologies, each of which holds different implications for how we should go about producing factual knowledge about world politics" (p. 32). The biggest challenge for Jackson lies in the fact that when confronted with ontological claims about the political world – whether they are about agents and structures, or states operating under conditions of realist interests, anarchy or global capitalism – there immediately arises the "problem of how possibly to know whether that claim is true, and (…) of

selecting the proper *methods* to use in evaluating the claim [original emphasis]" (Chernoff, 2009, p.391 cited in Jackson, 2011). I shall return to a detailed discussion of the 'ontology vs. mechanisms' question in Chapter 4, and selectively also in Chapter 3. Suffice to note here that this divide stems from deep philosophical and epistemological differences between the two camps, and that the PBP is trying to defend the merits of a 'mechanism' based approach based on insights on the workings of the social human brain.

In a similar vein to Jackson, political scientists Gary King, Robert O. Keohane and Sidney Verba propose in *Designing Social Inquiry: Scientific Inference in Qualitative Research* (1994) that social scientists should give up "parsimony" (ontological claims about the composition of the world) for "leverage" (a principle of hypothesis-building based on universal principles). The point of "leverage" is similar to that of the "bottom-up" approach discussed earlier: how can we theorize from the brain-level "upwards" to more complex phenomena such as political attitudes and behavior, without compromising the integrity of each level of analysis? By choosing "leverage" over "parsimony", King et al. argue for a social science that does not presume properties and characteristics about the political world and its inherent actors and structures, instead advocating for the construction of testable hypotheses through methodologies that are built on more universally accessible standards.

Critics of the "leverage" approach point out that committing to a single scientific methodology runs the danger of intellectual narrowness (Wight, 2006) and that the existence of universal standards cannot be defended epistemologically (Jackson, 2011). However, based on the internal discourse within social neuroscience outlined earlier in

this chapter, these fears might seem unfounded, at least if applied to the field of social neuroscience (also, King et al.'s idea of science is focused on political science rather than natural science). This chapter showed that social and political neuroscientists are well aware of the pitfalls behind extreme reductionism and determinism, and usually try to avoid drawing conclusions of this kind. Instead, they largely commit to the paradigm of multilevel integrative analysis, which acknowledges the multiple neural, genetic, epigenetic, endocrine, physiological, social and political factors and dynamics at play in determining the outcome of brain imaging data.

Further to this, social and political neuroscientists often have a differentiated and complex conception of "universality", usually resisting attempts at making sweeping generalizations for the larger social world. In fact, this chapter tried to show that self-critical and self-correcting reflexes and mechanisms are in-built within neuroscientific experimental processes, as well as their subsequent critical examination by fellow scientists in peer-to-peer reviewed research journals. Political scientists and philosophers who are wary of neuroscience's potential for reductionist universality need to acknowledge that they might be painting their image of neuroscientists and their ultimate aims with too broad a brush, to the point of constructing a conceptual and methodological enemy who might in fact turn out to be a straw man (at least in some parts).

The political scientists who were quoted earlier in the chapter and who critically engage with political neuroscience advocate unanimously that in order to overcome methodological shortcomings in neuroscientific approaches, what we need is a joint and interdisciplinary effort where political science can help refine the concepts and hypotheses that are tested experimentally through brain scanning methods, as well as

specify with greater clarity how exactly certain brain insights might apply to complex political scenarios. This is a very different approach to the "reductionist desire [by certain political scientists] of planting the findings of genetics, physiology, and neuroscience directly into the field of politics", as Maurizio Meloni points out in his critique of replacing a plurality of political frameworks with the intellectual authority of neuroscience (Meloni, 2012, p.30). If both sides can agree that the other side is indispensable in generating compelling political neuroscience data – as well as constructing viable neuropolitical theories about political perception, behavior and self – then there is no defensible reason why political scientists or political theorists should reject neuroscience completely.

Robert H. Blank (2014), in his edited volume on the current relationship between the life sciences and political science tries to remind political scientists that political science as a field has always been a "net-importer" of methodologies borrowed from statistics, economics and other disciplines – wondering why political science struggles so much to embrace or even just critically engage methods from the life sciences, including neuroscience. He warns that the current tendency within political science to center debates on a single approach or method for all social research is not tenable and is in need of critical self-examination and diversification (Blank, 2014, p.240; also see McDermott, 2011; Wiegele, 1979).

Blank's accusation might be too sweeping and somewhat unfair, given that current developments in political science methodology and field experimentation are trying to overcome the "single-method trap" by mixing quantitative and qualitative techniques, as well as advancing the scope of field experiments (Druckman, Green,

Kuklinski & Lupia, 2006; Humphreys & Jacobs, 2015; Humphreys & Weinstein, 2009). However, he manages to point to a more deep-seated resistance within political science that perhaps cannot be entirely be overcome through the mixing, refining and expanding of existing methods alone: namely, a resistance amongst a majority of political scientists to reconsider and reconceptualize very fundamental assumptions about human nature and motivations in social settings, as well as the cognitive abilities of political actors to interact and cooperate with others, based on breakthrough insights from the life sciences, in particular the neuroscience of the social human brain. The three following chapters try to show that this fundamental reconsideration is of utmost necessity.

3

3. The Rationally Dehumanizing Brain: Outlining Basic Neuropolitical Duties of Public Representatives in a Liberal Democracy

3.1. Social Cognition's Role in Political Inclusion and Exclusion

In this chapter, I attempt to construct a normative, interdisciplinary political theory about exclusion that draws on empirical evidence from social neuroscience and psychology. I argue that in order to grasp a major dimension in the identity conflicts we witness today, as well as in order to put forward truly effective normative guidelines for solving these conflicts, we need to understand the *social cognition network* implicated in including and excluding others, as well as more particularly, what kind of psychological and neurological mechanisms are at play when the worst kind of exclusion – dehumanization of fellow human beings – takes place.

The following questions drive this inquiry: what if *zoon politikon* is cognitively ill-equipped to handle one of the defining features of our modern world – hyperdiverse and hypermobile societies, in which distinct and often opposite identities abound, and are

constantly in flux[16]? Put differently, what if experiencing an excluding reaction to someone – to the point of shunning them from our idea of humanity – is in fact an integral part of how our brains make sense of their social surroundings? Furthermore, what if this cognitive disposition of ours is at odds with the norms of inclusion and tolerance required in today's liberal democracies, where rapidly diversifying communities are expected to co-exist peacefully with one another and cooperate on complex social issues? What if humanizing others is a distinct cognitive activity which actually requires conscious effort and work, but that political theorists so far have not paid enough attention to?

Presently, existing political theories on multiculturalism and the politics of difference make various assumptions about the extent to which human beings are able to exercise tolerance and inclusive attitudes towards others. Whilst some of them are more skeptical towards our innate capabilities of inclusiveness (Barry, 2001; Walzer, 2005), many of them assume that the reason for a lack of social cooperation between members of different groups is a belief in misguided values (Kymlicka 1995; Nussbaum 2002a; Patten, 2014) or a lack of relevant information on the actual benefits of inclusion (Kymlicka, 1998).

I argue in this chapter that many of these assumptions about our cognitive capacities for inclusion are misguided, due to being based on what I loosely term the "rationalist tradition". The actual picture that is emerging about our social brain is a

[16] One could object here that the phenomena of hyperdiversity and hypermobility took place in pre-modern societies as well, such as for example during the Ottoman Empire or Tang Dynasty China. However, sociologists (Beck 2007; Castells, 1997) contend that only in modern societies are identity struggles pushed fully into the public and global arena, and acknowledged in their own right.

challenging and deeply concerning one: empathy for and humanization of others is a limited resource (Decety and Svetlova, 2012; Vaish and Warneken, 2012; Fiske and Harris, 2006); failures of one's in-group are experienced as painful whereas failures of a rival outgroup give pleasure (Cikara, Botvinick, & Fiske, 2011); stereotyping individuals along their perceived status and human warmth is a universal phenomenon (Fiske, Cuddy, Glick, & Xu, 2002; Cuddy et al., 2009); dehumanized perception of others is an everyday, rapid and often passive process that once served evolutionary purposes in the context of living in groups competing over limited resources (Lee & Harris, 2014; Pagel, 2012); and finally, neuroimaging studies shows that we are highly sensitive to physical markers in others such as skin color (Kaul, Ratner, & Van Bavel, 2014), and are inclined to prefer own-race faces to other-race faces (Golby, Gabrieli, Chiao, & Eberhardt, 2001).

This chapter *employs the PBP to access the dynamics at play in the politics of exclusion* – assuming that in order to arrive at a sustainable politics of inclusion, we need to pin down the cognitive mechanisms and neurobiological circuitries at play when we exclude, reject and dehumanize others. If we do not start from this premise, our expectations about the conditions under which people can engage in politics, as well as our normative demands of how different identity groups should live and cooperate together might at best be too optimistic, and at worst, be based on assumptions that are opposite to what our social brains can in fact accomplish.

This chapter focuses on dehumanized perception firstly because of how it disrupts a vital social brain function called *mentalizing* (which makes up the basis for more complex social emotions such as empathy and compassion) and secondly because of how its *rapid, automatic and often passive nature makes it a politically challenging cognitive*

capability in the context of building political cooperation and solidarity across different groups. I examine how and why the "rationalist tradition" overlooks the potentially disrupting effects of dehumanized perception, as well as how it underestimates the fundamental role dehumanization plays in basic human social cognition. Hence the main aim of this chapter is to show that dehumanization – as studied by social neuroscientists and psychologists – matters for political theorizing about cooperation and solidarity, in particular within social contract, multiculturalism and cosmopolitanism theories.

A second aim of this chapter is to examine who in particular, in the context of a liberal democratic society, needs to avoid dehumanization because of how it could disrupt their mentalizing ability. Although it would certainly be desirable if all members had awareness and competent command over their spontaneous denial of humanness to various groups and individuals, it would be both unjustifiable and unfeasible to demand this to the same extent from every single member in a liberal democracy. This chapter therefore specifies one group in particular that we can be justifiably demand to mentalize and avoid cognitively dehumanize others: *public representatives*. Public representatives such as judges, law enforcement officers, bureaucrats and elected politicians have a different cognitive responsibility to other members in society – unlike any other member in society, they have a *special duty to mentalize and cognitively represent a wide and diverse constituency*, partly because that is their job requirement and that is what got them elected to public office, but also, being in command of the dehumanization capabilities that are part of our "flexible social cognition" (Harris, 2017) allows public representatives to be more competent, efficient and fairer towards fellow members within the liberal democracy whom they are accountable to.

The cognitive responsibilities of public representatives further have to be couched in the context of political deliberation and power relations, which is why this chapter distinguishes between two kinds of politics: the *politics of responsive equality* and the *politics of unresponsive inequality*. In both cases, functioning mentalizing abilities and a lack of dehumanized perception are crucial for the success of deliberative democratic practices as well as the prevention of abuses of power in the context of law enforcement, for example.

3.2. Dehumanization as Part of Everyday "Rational" Cognition

Recently, researchers of dehumanization have reached the preliminary conclusion that "conceiving of others as less human may reflect a basic, relatively passive, cognitive-perceptual process" (Hodson, MacInnis, & Costello 2014, p.90) and that dehumanization is "rooted in ordinary social-cognitive processes" (Haslam, 2006, p.252). Only in the last decade has dehumanization been considered a subtle, everyday phenomenon that permeates all aspects of our social interactions. Previously, social psychologists studied dehumanization in the aftermath of WWII and the Holocaust as an unusual and extreme form of exclusion that would result in aggressive and excessive behavior (Bandura, Fromson, & Underwood, 1975; Staub, 2010). Dehumanization was associated with acts of extreme violence, such as war killings and genocide, and was thus expected to occur only in exceptional circumstances (though these acts were thought to be committed by ordinary individuals).

With the groundbreaking work of Leyens and his colleagues (2000) at the turn of the century, dehumanization was suddenly recognized as a sweeping phenomenon that we all engage in on an everyday basis. They described this subtle and often unconscious dehumanization process as *infrahumanization*, in which we only ascribe uniquely human

emotions ('secondary emotions') to people whom we consider to be part of our in-group. These complex secondary emotions (e.g. embarrassment or optimism) are denied to people from out-groups; instead, we are only able to ascribe to them basic, non-secondary emotions such as fear and pleasure, which we would also attribute to other animals. Moreover, instead of aggression, one behavioral result of dehumanization is neglect.

Previous studies on prejudice and exclusion showed how belonging to an in-group and the mental construction of out-groups is part of how we function socially (Brewer, 1999; Baumeister & Leary, 1995; Krill & Platek, 2009; Tajfel, 1981). Following Leyens' infrahumanization model, researchers examined more closely how we subtly deny humanness to others. One important contribution has been Haslam's (2006) two-category model, according to which we dehumanize others along the lines of animalistic or mechanistic dehumanization. People who are dehumanized *animalistically* are denied 'human uniqueness', in that such people are seen to lack refinement, morality, civility and rationality (e.g., stereotype of the coarse and uneducated Mexican immigrant); people who are dehumanized *mechanistically* are denied 'human nature', in that such people are considered to lack interpersonal warmth, agency, emotions and depth (the stereotype of the over-achieving and ruthless Asian or Jew, for example).

Although the categorization of dehumanized perception has important political implications for understanding when and how exclusion occurs, I am focusing here on the more implicit cognitive processes that *precede* the explicit categorization of those whom we deem to be less than human. For example, one body of research looks at dehumanization as a basic, largely implicit cognitive process (Harris & Fiske, 2011). By studying the phenomenon at the neurological level with fMRI techniques, Fiske and

Harris (2006, 2007) discovered that dehumanization severely compromises a socially vital brain ability called *mentalizing* or Theory of Mind (ToM) (Wagner, 2015). Mentalizing is defined as an individual's ability to recognize the concept of mental activity in others, i.e. the ability to grasp cognitively that people have mental lives, beliefs, dreams and desires (Wagner, 2015). The term was first used by primatologists in connection to the social behavior of chimpanzees (Premack & Woodruff, 1978) and has since occupied the research of social neuroscientists who try to map the human social brain (Saxe & Kanwisher, 2003; Young & Saxe, 2008). Mentalizing is now considered to develop in early infancy in humans (Wimmer & Perner, 1983) and to be a cross-culturally existent ability (Harris & Avis, 1991).

An absence of mentalizing is correlated with (and in some instances considered to cause) questionable moral judgments and utilitarian decisions in contexts that involve other individuals (Koenigs et al., 2007). In its clinically pathological manifestation, people with autism are unable to execute the ability to mentalize others (Frith, 2001), leading some researchers to conclude that "[any] psychopathology almost always involves disturbances of social reasoning and theory of mind" (Brüne and Brüne-Cohrs 2006, p. 451). The latest research into mentalizing considers mental state reasoning to be a "critical cognitive input for behavior explanation, action prediction, and moral evaluation" (Young and Waytz, 2013). The capacity for understanding other minds is now seen as a capacity that has allowed humans to operate effectively in large social groups (Karmiloff-Smith et al., 1995; Tomasello et al., 2003). Also, the ability to read other people's minds has now been firmly linked to core interpersonal and social emotions such as empathy and compassion.

The picture that emerges is one where mentalizing is an indispensable social cognitive capacity for any individual in any society or culture. It is therefore all the more

unsettling, going back to Harris and Fiske's dehumanization research, that dehumanization can severely compromise our mentalizing functions. Politically, mentalizing thus emerges as a crucial cognitive activity in the creation and maintenance of social cooperation. I will return again to Harris and Fiske's treatment of dehumanization as an implicit cognitive process in my critique of bounded rationality theory below.

3.3. Dehumanization Blindspots

3.3.1. Dehumanization Blindspots I: Rational Choice

Dehumanized perception, as understood by social neuroscience and psychology, poses a challenge for what I would loosely term the 'rationalist tradition' in politics. I argue that in this tradition the failure to humanize each other is not considered central as to why social conflicts occur. At least two kinds of rationalist traditions come to mind here. One is centered on the rational choice paradigm, in which, classically, we expect economic and political decision-making to conform to a utility-maximizing model. This model was subsequently questioned by the empirical work of Daniel Kahneman and Amos Tversky (1974), which highlighted the susceptibility of human reason to bias and inconsistency, and subsequently paved the path for a new subfield devoted to the study of "bounded rationality" (Simon, 1982).

Although bounded rationality assumes that human behavior and decision-making is intendedly rational, it concedes that most of the time, rationality is severely constrained by various factors present in the internal and external environment of social actors – such as emotions, lack of adequate information, and situational incentives and constraints (Jones, 1999). Some theorists of bounded rationality have singled out our "human cognitive architecture" as the cause for these instances of rational failure (Ibid., p.298). By now, the idea of bounded rationality is recognized to the point that most economic

and political scientists accept that cognitive constraints can affect judgment and behavior at a significant level.

One of the most recent publications on this by political economist Johnathan Bendor (Bendor, 2010) presents us, three decades after Kahneman and Tversky's initial work, with an expanded picture of the kinds of cognitive restraints bounded rationality theorists have managed to narrow down. "What properties", Bendor asks, "should be on any political scientist's short list of cognitive features to consider when constructing behaviorally plausible theories of political decision making?" (Ibid., p.16). He offers the following list of six essential properties:

1. Top-down processing: our perception is overwhelmed by an abundance of environmental information, hence we process information selectively based on larger schemas and mental constructs.
2. Conscious thinking and attention is 'serial' in nature and cannot be expected to happen continuously.
3. Humans process information more slowly than computers, due to the physiological limits of neuronal signal transmission vis-à-vis electrical circuits.
4. Humans are inferior in calculation compared to computers and calculators.
5. Memory is actively reconstructive, not photographic.
6. Short-term or working memory is limited in scope, affecting the accuracy of long-term memory and thus, information-processing in general.

Except for 4. (the mental activity of calculation), the listed cognitive processes are predominantly unconscious and fairly passive cognitive processes; this makes sense, since bounded rationalists want to underline the effect of unintended influences in decision-making. However, I argue that one crucial, unintended cognitive process is missing in this list: dehumanized perception.

What is important about Harris and Fiske's establishment of a link between dehumanization and mentalizing is their focus on the implicit *cognitive processes that precede explicit categorization of people into respective dehumanized categories*, as well

as how their research is able to isolate exactly which mental capacities are affected by a lack of humanization. Their work grew out of the stereotype content model (SCM) (Fiske, Cuddy, Glick, & Xu 2002), which posits that social stereotypes are commonly made up along two dimensions, warmth and competence, discovering that only with stereotyped groups that elicit a disgust reaction (considered low in warmth and competence) does dehumanization occur. The treatment of *dehumanization as an implicit cognitive-perceptual process* (that can only be detected with fMRI scan techniques, versus, say, survey methods, which chiefly gauge attitudinal and behavioral manifestations of dehumanization, and would therefore be unable to access these implicit processes) allows us to argue for the inclusion of cognitive dehumanization into Bendor's political scientist's short list of cognitive features to consider in the construction of behaviorally plausible theories of how politics and social cooperation work.

Dehumanization, just like the other cognitive features on Bendor's list, is such a pervasive, basic and yet predominantly unconscious mental process, that it is highly inconceivable that such a fundamental cognitive-perceptual process would not influence the decisions of political actors in some significant way. Yet the rational choice tradition has hitherto overlooked the importance that dehumanization plays in the dynamics of social cooperation. Therefore, in this chapter, I argue that in order to account for a major feature of decisions made in a social context, political scientists who adhere to the rational choice model, or its modified cousin – the bounded rationality model, have to take into account that

A. Human beings ***routinely dehumanize*** others and
B. that these cognitive processes are implicit, automatic and often passive (Lee & Harris 2014), and

C. that dehumanization is relevant politically because it severely compromises a core social cognitive function called ***mentalizing***, and that therefore
D. any theory of human behavior and social cooperation has to include the uncomfortable fact that dehumanization of others can be one of the major ***disrupting cognitive abilities that can threaten social cooperation in hyperdiverse and hypermobile societies***, where people are expected to be able to include a wide array of individuals into their idea of humanity and circle of moral concern.

It is important to realize here that the inclusion of Haslam's two-category dehumanization model into Bendor's short-list, for example, would not make sense from the bounded rationalists' viewpoint. The fact that we commonly dehumanize people into these two categories and deny them human uniqueness and human nature would be considered too content-specific for rational theorists and would probably be grouped under informational bias (point 1. in Bendor's list). Indeed, Bendor explains that most rational choice/bounded rationality models aspire to the "scaling principle of modeling" where

> "(…) what matters in a model is not so much how sophisticated the agents are assumed to be or how hard the problems are, but rather the difference between the two. Typically, real humans are more sophisticated than agents in bounded rationality models, but real problems are also harder, both are scaled down in models. As long as a model scales down both sides symmetrically, it may be plausible even though the agents in the model are quite dumb."
>
> (Bendor, 2010, p.15-16)

According to the scaling principle, dehumanization could then just be treated as one specific informational bias and not be included in Bendor's list as a more fundamental cognitive feature that affects decision-making and social cooperation in its own right.

This is where Harris and Fiske's insights and their employment of fMRI techniques bring about a turning point. By examining dehumanized perception one step *before* explicit dehumanized categorization takes place, they are able to single out

dehumanized perception as a distinct cognitive process in its own right. Only through fMRI methodology were they able to determine that the brain regions that were disabled during dehumanized perception, such as the medial prefrontal cortex (mFPC), were those regions usually active during *mentalizing*. Based on this, they were could conclude that the effect of everyday dehumanized perception is much more disabling of our *social cognition network* than previously assumed. This is why whereas different dehumanization categories could be treated and ignored as contingent factors in Bendor's model, dehumanized perception is so fundamental to human social cognition that it cannot be left out.

Politically, we can then argue that even if in our models we would simplify agents' cognitive capacities and biases to a very basic level, dehumanized perception and the way it affects mentalizing cannot be removed from any model of social cooperation, if we want to retain some degree of plausibility and accuracy about how decision-making and social cooperation takes place in the political world. This leads me to the second way in which the "rationalist tradition" can be understood, namely through the political ideologies of Marxism and liberalism.

3.3.2. Dehumanization Blindspots II: Liberalism

One of the defining traditions within liberalism is social contract theory. The social contract – traditionally understood as collectively binding due to the power of individual consent by all its members – is thought to have been revived in the 20th century by John Rawls, amongst others. In his seminal work *Theory of Justice* (Rawls, 1971), normative principles such as those surrounding justice and just distribution of

society's resources "are conceived as principles that would be chosen by rational persons" (Ibid., p.16), and are able to subsume a 'plurality' of claims and interests by various groups under the common umbrella of one basic social contract. Indeed, Rawls *Theory of Justice* can be seen as the attempt to theoretically construct the conditions necessary for people from diverse backgrounds to be able to cooperate together and more ambitiously even, agree on hefty normative questions such as who deserves how much of the economic, educational and political share of available social goods in a liberal democratic society.

In order to accomplish this and to address the problem of diverse and opposing interests that would inevitably obstruct such an endeavor, Rawls comes up with by now famous concepts such as the 'original position', 'veil of ignorance', 'reflective equilibrium', and others. With the 'original position' – which describes the unbiased and fair viewpoint people are supposed to adopt before engaging with their fellow citizens in a consensus-reaching process of finding the ultimate principles of justice everyone can agree on –, and the 'veil of ignorance' – which refers to the hypothetical ability of people to be able to ignore their particular position in society in order to make decisions about these fundamental principles, Rawls introduced a distinct cognitive model of the rational citizen.

Rawls' critics have spilled much ink on the question whether his idea of the rational citizen and later, of political 'reasonableness', makes sense at all (Gauthier, 1974; Scanlon, 1973; Sandel, 1982). Critics have questioned how realistic it is to assume that people are able to think about questions of social justice in a cognitive vacuum, such as the 'original position', in which people are expected to put aside for a moment their

particularistic interests; or decisions made under the 'veil of ignorance', which assumes certain cognitive biases can be consciously withheld. Rawls himself and his defenders (Freeman, 2007) have countered to this that the theories in *Theory of Justice*, like many theories throughout the history of political philosophy, were intended as a non-descriptive thought-experiment and were part of what Rawls called 'ideational theory'. In other words, Rawls, just like Bendor in his 'scaling principle of modeling', tries to slim down the messiness of social reality to an idealized model of our cognitive dispositions *before* we enter social cooperation with others.

In this context, I claim that dehumanized perception is absent in Rawls' social contract model and I argue that it needs to be included, in order to express a *plausible* theory of individuals as social and rational beings. Again, just as earlier with bounded rationality theory, one could counter to this that dehumanizing could simply fall under the 'veil of ignorance' and that it is exactly this kind of cognitive bias that Rawls asks us to put aside in his thought-experiment. However, I treat dehumanized perception as such a fundamental cognitive ability that affects how we are able to mentalize those around us, and, which (as research at the neurological level showed) *precedes* opinions and judgments we might form about people, that therefore I claim that even if we fully recognize Rawls's project as an abstract thought-experiment, we still have to consider, even under the most ideational and abstract circumstances, what effect dehumanized perception could have on people who are supposed to cooperate with each other. In other words, the way I argue that Rawls' theory of social cooperation is deeply problematic is that he did not even consider it possible that the ideal rational person could fail to

cooperate with others on the grounds that she did not think them worthy enough as human beings.

Whilst Rawls accounted for all other ways in which political actors could possibly reject each other's viewpoint – namely along the lines of differing socio-economic theories of justice (the backdrop of the Cold War and the ideological battle between Communism and liberalism permeates *Theory of Justice*) and of ontological battles about existential truths such as in the realm of religion, which he addresses in his second seminal work *Political Liberalism* (published at the end of the Cold War and in the wake of the rise of identity politics) – he did not think that on a much more basic level of humanization, people's cognitive endowments could fail them and bar them from even the most plain type of cooperation with others.

Here is an additional example of the humanization problem in Rawls' theory. In *Theory of Justice*, Rawls makes what he calls a 'special assumption', which is that "a rational individual does not suffer from envy" (Rawls, 1971, p.143). It is peculiar that of all the caprices of human psychology, Rawls chooses envy. He explains that this is because "envy makes everyone worse off, [it is] collectively disadvantageous" (Ibid., p.144). A non-envious person is "not ready to accept a loss for himself if only others have less as well" (Ibid., p.143). In other words, what Rawls worries about most is that people will resort to vengeful intentions and actions if they let envy conquer their minds, disturbing the equilibrium of the entire social contract. What Rawls does not consider is that if I envy someone, I might have already humanized them to a large extent, actually so much so that I want to possess what they own, consume what they consume, do whatever activities they engage in – to actually wanting to *be* whom they are!

In the "stereotype content model" (SCM), where stereotypes are considered to be of mixed nature, envy falls in the dual categories of low warmth/high competence, meaning that we perceive people we envy as highly capable and competent, but attribute few empathic and warm characteristics to them. In other words, we might admire people whom we envy but we do not find them very likeable. In the US, for example, social groups that fall into the envy category are Asians, Jews, and the rich (Fiske, Cuddy, Glick, & Xu, 2002). Stereotypes are therefore tricky to address and remain persistent because like envy, they contain mixed categories of warmth and competence, i.e. people who are stereotyped this way might be told that the stereotyping is meant as a compliment, since envied people, for example, are viewed as highly competent. A recent international study that tested correlations between income inequality and the presence of the SCM model's mixed stereotypes detected that the higher the income inequality, the higher the presence of ambivalent stereotypes (Durante et al., 2013). More equal societies, on the other hand, were less tolerant of mixed stereotypes. In that sense, Rawls' assessment of envy's destabilizing effect on a just society was correct.

However, in terms of dehumanization, Harris and Fiske's research linked dehumanization to disgust, in that amongst the main four stereotype categories of envy, pride, pity and disgust, only disgust groups were seen as less-than-human. This let them conclude that only with disgust do we activate brain parts that we usually reserve for the perception of objects, not humans. In the US, social groups in the disgust category include welfare recipients, the homeless, Arabs and poor Blacks. Further, the significant *behavioral* difference arising from disgust vs. envy is that we will direct aggressive

actions towards people whom we envy, whereas we simply neglect people for whom we feel disgust (Cuddy, Fiske, & Glick, 2007).

This makes sense, since we do not bother about objects the way we care about people; therefore, people whom we view as less-than-human are exactly that – excluded from our vision of who belongs to society and treated like invisible objects no one cares about. Linking this back to political thinking on racism, for example, Ralph Ellison's *Invisible Man* (Ellison, 1997) and his definition of blackness as invisibility come to mind, as well as Martin Luther King's statement that he was "fighting the forever degenerating sense of nobodiness" (Cited in Mills 1997, p.112). At least the envied people are given a visible seat at the table, even if we might want to snatch it from them!

The treatment that dehumanized groups and individuals receive in the form of *neglect* is therefore the worst that can happen to a party who wants to join the social contract. This is why I argue that Rawls, and the liberal tradition that surrounds him, misses the fundamental cognitive condition necessary for inclusion and equal treatment, which is humanization of the other. As mentioned above, the problem in Rawls is that he thought people would reject each other mainly along the lines of differing socio-economic visions of society or, as he went on to include in his late life, clashing comprehensive doctrines such as religious ones. He simply did not think it possible that the process of agreeing to principles of justice between differing parties could fail because some are unable to perceive the other as human in the first place.

Charles Mills, critical philosopher of race, criticized Rawls and the Western social contract tradition in his *The Racial Contract* (Mills, 1997) for the existence of a 'parallel discursive universe' (p.131) in which 'white moral cognitive dysfunction' (p.95) denied

people of color the full 'moral state of personhood' (p.118). Although Mills' critique is compelling, it is not fully satisfying because he lacks empirical psychological and neuroscientific evidence of exclusion and dehumanization to make his case: Mills can point out the dehumanizing instances within Western political thought, but he is unable to provide further explanation about what exactly is going on in the brain of the excluder when this happens. In other words, Mills has lacks insight about the exact mechanisms implicated in cognitive dehumanization and in which way exactly dehumanization would affect social brain functions such as empathy or moral judgments, for example. This, in turn, makes his otherwise powerful objections to social contract theory's White bias susceptible to the criticism of a new generation of philosophers of race, who are attempting to backtrack from Mills's more radical claims by offering alternative readings of White social contract theorists[17].

Thus what the brain science can offer political theorists is precisely an informed understanding of this cognitive mechanism and hard evidence for how dehumanization affects vital social brain functions such as mentalizing. Disagreeing with Rawls through this perspective then is not simply an aesthetic or moral divergence of political opinion but becomes an informed disagreement based on the recent brain insights into what it takes to view others as full human beings. It effectively suggests that Rawls' standard of 'reasonableness' for citizens requires the added category of humanization in order to make sense even just as ideational theory[18].

[17] See Brandon Terry's (2013) recent critique of Mills, based on an analysis of new archival material of Rawls's personal reflections on the civil rights movement.
[18] For a more standard critique of Rawlsian resonableness, see Jeremy Waldron's (2003) argument about the limits of Kantianism and Rawlsian reasonableness in the face of clashing views of the good life in diverse societies.

In other words, whereas other biases can be treated as contingent factors and therefore be justifiably ignored in Rawls' ideational model, I argue that *dehumanized perception is so fundamental to our cognitive make-up* that it has to be included even in the most abstract, ideational thought-experiment of social cooperation. In this sense, what the PBP can offer us is not just a more substantiated critique of White political theory's dehumanization blindspot, but even more importantly, it can contribute to a constructive, more informed reconstruction of liberal theory's norms and objectives for social cooperation in a hyperdiverse world.

3.3.3. Dehumanization Blindspots III: Marxism

The historical backdrop of slavery, colonialism, segregation, gender inequality and the general social inequities loomed at every point in history when major Western political works were written – from Aristotle to Augustine, to Machiavelli, Hobbes, Locke, Kant, Marx and far into the 20th century. Political thinkers were not only passive witnesses who inevitably internalized some of the dehumanizing social dynamics of their time, but in many instances, they were also active players within and defenders of dehumanizing systems, exemplified by John Locke's involvement in the slave trade and ideological justification of slavery (Bernasconi & Mann, 2005) or Immanuel Kant's belief, upheld until the end of his life, that blacks were inferior human beings (Gowans, 2001; Louden, 2011). Although it has been suggested that Kant changed his views on colonialism as he got older, recent scholarship suggests that even though that was very likely the case, he (much like Thomas Jefferson) did not change his views on race and blacks as sub-humans, who were not considered to be part of his circle of human concern

(Louden, 2011, p.134)[19]. In other words, although Kant was able to reject colonialism on ideological and ethical grounds, he might not have been able to humanize and mentalize the actual people who were suffering within that unjust system.

This phenomenon of the peculiar coexistence between the acknowledgment of faulty ideological systems and yet the *inability to humanize those who are suppressed within those very systems* is nowhere better manifested than in Marxism's dehumanization blindspot. For Marxists, the phenomenon that "the hungry don't steal and (…) the majority of those who are exploited don't strike" can be attributed to the idea of "false consciousness" (Cited in Rosen 1996, p.1). Although Marx and Engels never explicitly named this idea in their public writings[20], it permeates much of their understanding of the psychological disposition of human beings and the forces behind historical change.

"False consciousness" is a way of explaining why a group of people – usually the suppressed in society – comply with and obey their oppressors even though it goes clearly against their own interests. 20th century Marxists such as Georg Lukács, who coined the term in his seminal *History and Class Consciousness* (Lukács, 1971 [1920]), Herbert Marcuse (1964) and Theodor Adorno (1974) applied the idea to understanding why societies would stagnate in progressing towards a post-capitalist stage of historical development, which Marx predicted would inevitably take place, due to the inherent contradictions within capitalism. False consciousness offered a compelling explanation: the hungry don't steal and the oppressed don't rebel because they have not understood

[19] For a comprehensive discussion of Kant's relationship to colonialism see Flikschuh & Ypi, 2014; see also Kleingeld, 2012.
[20] Engels spoke of "false consciousness" only once in a private letter to Franz Mehring in 1893.

their inferior role in society yet – the belief that this realization will inevitably take place in the minds of the oppressed underlies the fundamental optimism of the Marxist vision of human psychology – but once they have rationalized their position, and unmasked the sinister exploitative forces that grip society in this backward state, they will rebel against the ruling class and bring about the next stage of history.

How is this related to dehumanization at the brain level? The example of false consciousness is helpful because it illustrates how Marxists believe that detrimental psychological dispositions can be overcome by reason and the exposition of truth alone. Indeed, Michael Rosen considers "false consciousness" an idea that represents a "rationalist conception of the good for human beings" (Rosen, 1996, p.274), which he traces back to a Western rationalist tradition spanning from Plato to Étienne de la Boetie, Adam Smith and the Enlightenment thinkers. Rosen is skeptical of this rationalist conception of human motivations and behavior, and although he acknowledges the voices of anti-rationalists such as Rousseau, Nietzsche, Walter Benjamin, Freud and others, he deems their critique of rationalism to "lack[s] the foundations that would be needed to constitute a truly effective alternative to the dominance of rationalism" (Ibid., p.274). Rosen presumably is referring here to the lack of an empirically grounded foundation that would have the necessary evidentiary and theoretical force to counter rationalism in its current form. It is exactly this lacking foundation that a neuropolitical, social cognition-based critique of the current rationalist tradition could provide.

If we apply Rosen's analysis of Marxist rationalism to dehumanization, the argument could go as follows: from a Marxist point of view, unequal and exploitative relationships – of which dehumanization is one – persist because both the exploited and

the exploiters (that is why even capitalists are considered to be capable of joining the Communist movement, once they come to understand the sinister workings of their actions) have not understood the underlying socio-economic mechanism and deeper moral truths of their relationships. The problem here for Marxists is not one of emotional motivation, but of rational epistemology. Once the actual truth is understood, the exploiters and most certainly the exploited will change course and redefine their political aims and sense of self.

Has the Marxist hope of overcoming detrimental psychological processes through reason been replaced by a more sensible and nuanced perspective today? This does not seem to be the case. Jane H. Hill, a linguistic anthropologists studying the everyday language of White racism, reports that whenever she discusses the pervasiveness of racism in the US across university campuses and in front of various nationwide audiences, fellow White people (including her academic colleagues) will react with defensiveness and anger, and usually tell her the story of what she calls the "folk theory of racism" (Hill, 2008, p.5). One part of this folk theory consists in believing that racism is a matter of individual beliefs, that racists are anachronisms and that ignorance will eventually be cured by education and well-being (Ibid., p.8). Defenders of the folk theory of racism cannot accept that prejudice and extreme exclusion might be an integral way of how human beings (including themselves) function and that dehumanization of others *cannot* be cured through an act of will and reason alone.

Indeed, to show how this folk theory of racism extends in more sophisticated political debates in the field of multiculturalism, it is worth referring to Anthony Appiah here, a prominent British political theorist of Ghanaian descent, who once responded to

the well-meaning suggestion by fellow liberal multiculturalist colleagues of overcoming White prejudice against African-Americans through educating White children early on in school about the merits of African culture and civilization, with the exasperated statement: "It is not black culture that the racist disdains, but *blacks* [emphasis added]" (Appiah 1996, p.26). This is a powerful example of how the rationalist tradition thus keeps exerting its powerful influence beyond Marxist borders into liberal multiculturalist theories. By claiming that dehumanization and extreme cognitive exclusion is a problem of the unreasonable or uneducated mind, instead of our emotive and cognitive brain limitations, multiculturalism theory is echoing the folk theory of racism, denying effectively that dehumanization is a disturbingly deep-reaching phenomenon and failing to acknowledge that racial exclusion is based on a fundamental social cognitive ability we all carry within us.

The dehumanization research of the last decade, as well as the much older field of prejudice study within social psychology, has been able to discover certain conditions that might lessen dehumanization of others and increase humanization of out-groups and the willingness to cooperate with them. The famous contact hypothesis, first posited by Gordon Allport (1954), states that social contact between out-groups can reduce prejudice, and has been successfully retested and refined since (Dovidio, Eller, & Hewstone 2011; Pettigrew, 1997). However, some political scientists counter to this that the contact hypothesis fails to account for the causes behind recent genocides such as in Bosnia and Rwanda, where ethnic and religious groups lived peacefully and in close contact with each other for many centuries. Here, social neuroscientists who study dehumanized perception at the neurological level might be able to offer a more nuanced

picture. They found that the tricky aspect about dehumanized perception is that it can be turned on and off fairly easily and often without us noticing; the nature of in-group categorization and dehumanization of others is automatic and spontaneous (Lee & Harris, 2014).

We further know that dehumanization is linked to mentalizing and mental state attribution of others, in that we fail to attribute mental states to those we dehumanize. One way therefore to counter dehumanization could be to actively and consciously mentalize social targets that we know are vulnerable to extreme exclusion in society. Although this conscious mentalizing might not make a big difference in peaceful times – in that there is little conflict between different groups in those moments – having exercised my mentalizing abilities and consciously humanized these groups might have a significant effect in moments of conflict, where the competition over resources and the situation of civil war puts great pressure on decision-making. This could potentially have important implications for IR theory, where scholars have long been aware of the precarious cognitive situations under which decision-making takes place (Jervis, 1976) and have recently payed increased attention to the role played by cognitive and emotional factors in conflict situations and negotiations (Crawford, 2009; Petersen, 2011).

Furthermore, researchers of prejudice have found that playing simulated cooperative games can reduce cognitive biases towards outgroups of another ethnicity (Sheng & Han, 2012), through which self-identification with out-group members is facilitated. In addition, the cognitive activity of individuating other people, i.e. imagining people as distinct individuals, can also contribute to humanization of the other (Swencionis & Fiske, 2014). Latest findings in the nascent field of cultural neuroscience

indicate that empathic abilities might even be uniquely shaped by culturally distinct construals of the self (Sheng, Han, & Han, 2015; Zuo & Han, 2013).

In sum, the ongoing research into prejudice reduction and rehumanization suggests that in order to tackle exclusion, we cannot rely primarily on our reasoning capacities but have to include emotive and mentalizing capacities if we want to humanize, de-objectify and individuate others. Irrespective of whether we are talking about ideational or non-ideational theories, I claim in this chapter that this is what the rationalist tradition needs to pay attention to, if it wants to work with plausible assumptions about human cognitive capacities and behavior – as well as have a reasonable idea of where rationalist politics can begin.

3.3.4. Dehumanization Blindspots IV: Human Rights

In another strand of the rationalist tradition – the human rights discourse – it is not clear whether an explicit awareness about the cognitive challenge of humanizing others, as distinct from normative or ideological concerns, is present. A recent study on the effect of human rights education on the mindset of Indian police officers looked at the extent to which police officers were willing to adopt human rights norms and whether it changed their view on torture, after they had attended a two-year long human rights course (Wahl 2013; 2017). Torture and extrajudicial violence by police in India is pervasive to the point of routine (HR Watch Report 2009; Commonwealth Human Rights Initiative 2011). India's relationship towards the UN Convention against Torture is ambivalent; it signed the Convention in 1997 but has not ratified it yet. Similarly, domestic legislation against torture has stalled in recent years, generating much criticism from human rights activists.

India is an interesting case here because of its ethnic and religious diversity within a non-Western democracy.

In particular, the study looked at whether police officers could reconcile their local beliefs about torture with those universal values learnt in their two-year course. The findings were surprising in two ways: contrary to the accepted view in the literature on the incompatibility of local beliefs with Western human rights values, the Indian police officers found it relatively easy to incorporate human rights beliefs into their local religious and moral norms of kindness and respect towards others. However, they still reserved the right to torture certain people whom they did not consider worthy enough of humane treatment.

What this case underlines is the point I made earlier in connection to Marxism's dehumanization blind spot, which is that the denunciation of faulty ideological systems and yet the inability to humanize those who are suppressed within those very systems often coincides. After the two-year human rights course, the police officers were able to use Western human rights language and concepts to denounce the use of torture, yet because they still cognitively dehumanized certain groups in society, they were unable to extend human rights protection to these people. No matter how many human rights values those Indian police officers would be willing to adopt, the use of torture will not cease until they are able to fully humanize all of their constituent groups.

3.4. Political Contextualization of the PBP

3.4.1. The Politics of Responsive Equality

There exists a lack of clarity about which kind of politics we are referring to when we debate the exclusion, in the sense that dehumanized perception occurs and matters in situations ranging from representative democratic debates in Congress to election campaigns, affirmative action, multicultural policies, ethnic violence, police violence, state-led sterilization of the urban poor, and the implementation of human rights – the list goes on. In any context where people from different groups face each other, or even just where two different individuals engage in a basic socio-political transaction or relationship, it matters whether these people are able to view each other as fully human, and as a consequence are able to attribute mental states to each other. This makes the study of dehumanization in political contexts both very important and very confusing. I believe that one major reason for the confusion is that we are not clear about which type of politics we refer to in a given context, and propose that there are two main types of politics that are at stake here: the *politics of responsive equality* and the *politics of unresponsive inequality*.

In the *politics of responsive equality*, politics happens predominantly in institutionalized and civil contexts. People there have the time to debate with each other, listen to each other's arguments, sit down for negotiations and deliberate their political future. In this context, dehumanization matters because it can severely cloud someone's judgment of the situation, and most important, can potentially make *political persuasion based on public reason impossible*. Some adherents of the rationalist tradition (though not Marxists) fall into this category, such as rational choice theorists, deliberative democratic theorists such as Jürgen Habermas, liberal theorists such as Immanuel Kant, John Rawls or Brian Barry, and multicultural theorists such as Will Kymlicka. To put it

more bluntly, the kind of politics envisioned there does not involve guns, violence or bloodshed, but treats political actors as reasonable actors who can disagree with each other profoundly, but could potentially be swayed to an agreement based on an appeal to their shared public reason and through the process of political deliberation.

Though one might say that this vision of politics is too stylized and naïve, it is essential for any democratic polity – in order for us to locate politics in the procedural and deliberative institutions that we erected and maintain as democratic citizens, and for us to have an idealized society in mind that we can strive towards. It is, however, an incomplete picture of politics. It is therefore not surprising that this kind politics, where everyone is naturally assumed to have an equal say and assumed to be given an equal place to say it without fear and repercussions, is criticized for its aloofness and irreality by African-American political philosophers (Mills, 1997; Shelby, 2005) and by critical political theorists, who base their idea of politics on a Marxist view of politics as a power struggle and inherent exploitation.

Yet this chapter argues that dehumanization matters for the politics of responsive equality, and that the pursuing of this kind of politics, even if it works with an idealized model of society, is necessary in order for us to think and debate about how we would want to live with each other, under conditions where we would be able to humanize each other most of the time. This is why I argue that this kind of political model has to take into account everyday dehumanized perception and recognize mentalizing as a crucial social cognitive ability in order to work with plausible models of human behavior and have convincing expectations about how social cooperation works.

In the context of the multiculturalism debate and whether liberal societies should reserve a special status for minority rights, Brian Barry in his book *Culture and Equality* (Barry, 2001) rejects the justification for special minority accommodation and argues that the "politics of difference" should be replaced by what he calls the "politics of solidarity", which consists in the neutral treatment of minority out-groups and universal equality for all. Barry's view of society is severely limited by his own epistemological and biographical narrowness[21] and lack of citations of empirical studies on stereotyping and the status of minorities in the West. However, his view is worth mentioning here because it represents a widespread belief amongst liberal political theorists: the idea that equal treatment of all will happen simply by persuading people that this is what they should value. This however, I argue, paints a distorted picture of the politics of responsive equality, in which equality is understood solely as a rational value or ideology, but not as a cognitive-perceptual ability of the brain that can only be reached once we have tackled our everyday dehumanized perception of others.

3.4.2. The Politics of Unresponsive Inequality

In the *politics of unresponsive inequality*, there exists no leisurely time to stand at a door step and engage in a conversation with a hitherto dehumanized out-group, allowing an engaging conversation to induce mentalizing and reduce dehumanized perception. In this kind of politics, an adult, sometimes even a child, gets shot by the

[21] Barry suggests that Asians and Jews do not face discrimination in US society because – based on his own narrow experience of having lived on three university campuses and briefly around Chicago as a British White male – Barry's reasoning for this was that Asians in his perception were successful and Jews able to get kosher food and observe their religious practices in the Chicago area, see 306ff.

police even before they had the chance to defend themselves with words, and are given no platform to reduce the prejudice that is directed towards them through an engaging conversation. The recent and ongoing police violence against young African-American men inevitably comes to mind here, such as the case of Tamir Rice, the 12-year old boy in Cleveland who was shot by a police officer before the officer had even confronted him, only knowing that the suspect was black[22].

In these instances, the power hierarchies – both in terms of physical and socio-political power – are so unequal that there is simply no time for the kind of prejudice reduction that could be successfully implemented in the realm of the politics of responsive equality. Yet the experiments studying prejudice and dehumanization reduction assume that the dehumanizing individual will lend their time and attention to reducing their own cognitive biases – in fact, time plays such a crucial role in prejudice reduction, but is a limited resource in the politics of unresponsive inequality. In this second kind of politics, those at the receiving end of this unequal relationship feel that politics is far removed from a deliberative process or situations where dehumanized perception could be challenged in a meaningful way.

The significance of the lack of humanization of social targets in the cases of recent US police violence is signified through the slogans that protesters upheld, such as 'Black Lives Matter', as well as by the dehumanization research itself. Waytz and his colleagues (Waytz, Hoffman & Trawalter, 2013) found out that White Americans commonly "superhumanize" African-Americans, ascribing to them supernatural and

[22] https://www.washingtonpost.com/news/post-nation/wp/2014/11/26/officials-release-video-names-in-fatal-police-shooting-of-12-year-old-cleveland-boy/?utm_term=.2808233c994a, accessed February 3, 2017.

magical mental and physical qualities, which in turn leads to denial of pain of Black targets. This phenomenon could in fact be observed in the Michael Brown shooting in Ferguson, where the police officer was quoted in court of saying that Brown appeared to him like an incredible 'hulk' of supernatural physical height and that his face looked 'demonic'[23]. Because the police officer effectively dehumanized Brown by 'superhumanizing' him, this might have led to the excessive violence directed against him, since the officer felt that Brown would not feel pain the same way as actual humans do. Moreover, since dehumanized perception is known to happen only 100 *ms* after stimulus presentation (Willis & Todorov, 2006; Harris & Fiske, 2009) there is little left to do for these social targets of police violence to dissipate the bias directed against them, even if they wanted to engage in deliberation or meaningful communication with their perpetrators.

The "Black Lives Matter" movement (Taylor, 2016) recently became a focal point in the debate between Republican candidates for the 2016 presidential election, with many criticizing the movement's slogan itself. Kentucky senator Rand Paul for example suggested changing the slogan to "All Lives Matter", whereas Ben Carson, the only African-American Republican nominee, stated that the slogan is a distraction from real political problems facing the Black community[24]. With the brain evidence pointing to the way dehumanization shuts down crucial cognition abilities such as mentalizing and

[23] http://www.theguardian.com/us-news/2014/nov/25/darren-wilson-testimony-ferguson-michael-brown, accessed August 30, 2015. For Darren Wilson's complete testimony see State of Missouri v. Darren Wilson, Grand Jury Vol V, Sept 16, 2014 at http://www.documentcloud.org/documents/1371222-wilson-testimony.html, accessed August 30, 2015.

[24] http://www.theguardian.com/us-news/2015/sep/04/republicans-black-lives-matter-scapegoating-movement, accessed September 5, 2015.

empathic concern, as well as the way how Blacks are singled out in U.S. society for "superhumanization", the poignancy and necessity behind the idea of "Black Lives Matter" can be defended more powerfully. In turn, the dehumanization at the brain level emerges once again as a tangible and biologically grounded phenomenon with serious political consequences for the less powerful in the politics of unresponsive inequality.

We can thus generalize the political argument about mentalizing as follows: if dehumanized impressions are formed so quickly and if there are situations in which the politics of responsive equality stands no chance, then the conscious mentalizing of social out-groups by those who have disproportionately more power is an essential preventative measure against violence driven by dehumanization. Practically, this could mean programs in which police officers are routinely prompted to perform cognitive mentalizing exercises about vulnerable out-groups in their constituency, and where thus the conscious mentalizing of social targets is considered central in preventing the escalation of violence against them[25].

Racially biased police violence is not the only place where the politics of unresponsive inequality takes place – numerous other social groups are denied the cognitive attention, time commitment and face-to-face interaction necessary for mentalizing and re-humanizing them.

[25] Preliminary research on the cognitive and affective burdens of humanizing others suggests that excessive pre-mentalizing of an out-group might actually lead to increased dehumanization (Cameron, Harris, & Payne, 2016). This is because humanizing someone can be painful for us, so instead we choose to avoid engaging with them in the first place. Therefore an alternative way to moderate one's dehumanized perception might be to practice mindfulness whilst dehumanization occurs in a controlled setting; this way one learns to recognize and moderate everyday dehumanized perception in a sustainable and integrative way.

From prisoners, to single parents living on welfare, to asylum seekers and beyond – the structural inequities in society do not encourage our political, economic and intellectual (including academics) elites to communicate with or encounter these groups face-to-face on a daily basis. Thus one criticism that one could level against the laboratory settings in which social neuroscience and psychology experiments are conducted is that they chiefly mimic the politics of responsive equality, where people are required to devote cognitive resources and time to each other through institutionalized deliberative democratic processes.

Yet without incorporating the realities and dynamics of unresponsive inequality, where the scarcity of time and cognitive attention is pivotal in deciding life-and-death outcomes, laboratory studies on prejudice reduction are insufficient to capture the complete range of politics where cognitive exclusion and dehumanization can play a crucial role in terms of equal treatment, democratic deliberation, and political cooperation. In both kinds of politics, the responsibility of public representatives to be able to mentalize their diverse constituents and to overcome (or at least be aware of) the way how their brains dehumanize others routinely and spontaneously seems vital for the basic functioning and implementing of fair and equal treatment within a liberal democracy.

3.5. A Neuropolitical Opportunity for Critical Theory

In critical political theory there exists an acute awareness about the destructive effects of the politics of unresponsive inequality and the need to address the injustices arising from it. In particular, in the study of the "politics of difference", scholars have

talked about the need for an "ethics of recognition" (Honneth, 1996) and "equal recognition" (Patten, 2014) of disadvantaged minority groups, demanding for them "equal social status" (Fraser, 2000). The problem however with these theories is that their arguments lack the empirical and objective force that the psychological and neuroscientific research into dehumanization can offer, specifically, in terms of the exact cognitive mechanisms that are involved in dehumanized perception. Although these critical theories often demand humanization of out-groups in some form, they do not explicitly name this demand and instead use an array of different, often more vague, terms, rendering it difficult to find a more unifying concept under which the phenomenon can be understood and discussed politically.

Most of all however, they are unable to argue that mentalizing is a crucial social cognitive ability that is compromised during dehumanized perception and that therefore mentalizing should be treated as a politically highly relevant mental ability. In this way, critical theorists of exclusion do not manage to come up with a specific cognitive mechanism with which to describe and understand social exclusion, dampening the persuasiveness and impact of their theories beyond their own field significantly. Yet critical theorists hold a unique and crucial position in the debate on the politics of exclusion. No other theoretical approach currently focuses on the politics of unresponsive inequality with the same concern and analytical sophistication as critical theory does. One aim of this chapter was to show that in order to build a meaningful neuropolitical theory of exclusion, neuroscientific laboratory results cannot be simply applied to the rationalist individual constructed by the liberal White canon that most of Western political philosophy is built upon. Social neuroscience and psychology experiments have

been accused of being conducted in too sterile and artificial settings, yet this chapter argues that it is not artificiality that is the main problem here (in fact, many political processes are artificial), but that the majority of laboratory settings *exclude basic aspects of social structural inequalities* in their conception and set-up.

The neuropolitical opportunity for critical theorists lies in bringing to the table a rich tradition of analyzing and understanding these very structural inequities, and help incorporate this into the experimental design and the drawing of political conclusions in the current brain research on exclusion and re-humanization. For this to happen, however, critical theorists and the political Left have to overcome their deeply seated Marxist suspicion towards psychological explanations of social exclusion, by appreciating that the momentous unlocking of the 'black box' of our social brains, which is currently taking place at this extraordinary moment in the history of science, can potentially aid substantially in understanding the fundamental cognitive mechanism at play in social exclusion. In this sense, this chapter tried to show the brain science of exclusion can help strengthen political theories of structural inequities considerably, and lift them to another level of awareness about the cognitive conditions under which politics takes place.

For example, a potentially fruitful application of the PBP to critical theory could lie in the realm of political language. Mentalizing/Theory of Mind researchers have managed to show a strong link between language usage and mentalizing. For example, it is well established by now that linguistic ability relates to performance on tests that examine the presence of ToM (Astington & Jenkins, 1999; Astington & Baird, 2005). A speaker's ability to use so-called "mental state verbs" (e.g. think, know, want) when asked to describe someone else's state of mind is a reliable indicator for the presence of

ToM (Semin & Fiedler, 1988). If mentalizing language can be an indicator for the presence of a speaker's cognitive mentalizing activity, then language usage can also be an indicator of whether someone is actually humanizing the person they speak about.

This could offers political scientists a powerful analytical tool to analyze political speech for dehumanization, based on what kind of mentalizing words are used by the speaker in describing a certain out-group. With Marxism's concern about dismantling ideological superstructures, which are thought of as the linguistic and conceptual ideas passed down through elites, there is potential for a promising merging of these two ways of language analysis, and a chance for critical theory to pin down the neurolinguistic cognitive mechanisms at work in social exclusion, whilst being able to embed them in a materialistic critique of unequal and exploitative social structures.

3.6. Conclusion: The Neuropolitical Duties of Public Representatives

If excluding someone – to the point of shunning them from our idea of humanity – is in fact an integral part of how our brains make sense of their social surroundings, how can this fact be integrated in our thinking and the political structures and language of liberal democracies? What innovative ways can we outline, through this new insight into the cognitive mechanism of social exclusion (as well as the overcoming of exclusion and prejudice), of how to deal politically with a cognitive ability that once made sense evolutionarily, but is now misplaced in a globalizing world where social cooperation does and has to take place across cultural and identity boundaries?

This chapter attempted to make a first step towards formulating these questions out of the empirical evidence that is generated from the social neuroscience field, and

making them relevant for our political thinking about identity politics and beyond. The chief aim of this chapter was to offer a broadly-painted sketch of how political theorists could begin to apply the PBP and include dehumanized perception into a variety of social cooperation theories – spanning from social contract, multiculturalism and cosmopolitanism theories to deliberative and power relationships between constituents and public representatives in liberal democracies. What all of these theories have in common are their inherent "rationalist" assumptions about the human capacity for social inclusion.

This chapter tried to show that dehumanized perception, as understood and researched in the context of our social brain, needs to be treated seriously as a fundamental cognitive process that potentially undermines the way how we live and cooperate together politically. The everyday, rapid and automatic nature of dehumanized perception poses a particularly challenge to hyperdiverse and hypermobile societies where identities and allegiances are in constant flux, and yet where the peaceful coexistence and political cooperation between a wide array of socio-cultural groups is desired. In this context of hyperdiverse and hypermobile modern society, public representatives must shoulder a special cognitive responsibility: more than anyone else, they carry a distinct duty to be able to mentalize and cognitively represent a wide and diverse constituency, and make decisions on behalf of many mentalized positions, viewpoints and identities. Being in command of their mentalizing faculties and being aware of the troublesome political and social effects of everyday dehumanized perception, bias and prejudice – particularly in light of structural inequalities in terms of raw power between themselves and their constituents – seems to be crucial both for

public representatives' success at fulfilling the representative nature of their jobs, but also, to justify the positions of power they are holding on behalf of a democratic and diverse citizen body.

4

4. In Search of Dignity: Ontology vs. Neurocognitive Processes in Human Rights and Multiculturalism Debates

> Human dignity is one of those concepts that politicians, as well as virtually everyone else in political life, like to throw around, but that almost no one can either define or explain.
>
> Francis Fukuyama, *Our Posthuman Future*

4.1. Does Dignity Have to Exist?

The previous chapter explained why dehumanized perception is crucial in political settings where an awareness of the mental state of others is needed, in order for decisions and cooperation based on public reason to take place. I distinguished between two kinds of politics, the *politics of responsiveness* and the *politics of unresponsiveness* – arguing that in both cases, dehumanized perception is required to either maintain epistemological equality between different political interest groups or to prevent gross power inequalities (with potentially ensuing violence) between them. I assumed that dehumanization, or the denial of humanness to someone, would be deemed undesirable based on the political and moral premises liberal democratic societies are built on.

Even though this chapter maintains this point, I argue that within current theories of toleration, multiculturalism and human rights, there is little or insufficient weight placed on the cognitive challenge of humanizing someone in our everyday social and political lives. I argue that this stems from confusion about how humanization is achieved and about what makes us *dignified* human beings in someone else's eyes. My main argument here is that currently, the idea of dignity – and not humanization – lies at the heart of theories such as universal human rights, and that dignity there is imagined and pursued as an *a priori* ontological "essence" inherent in all human beings. Similarly, in other theories such as the politics of difference and multiculturalism, the main goal seems to be to acknowledge that individuals or groups who have different cultural or religious life pursuits and customs to the majority are nonetheless entitled – due to the inherent worthiness or dignity of the 'otherness' of their identities – to non-interference or even protection by political authorities. In all of these cases, it is the contestation of dignity,

not of humanization, which is seen as the ultimate battle ground for the overcoming of interpersonal-and group conflict and the protection of political rights.

Instead of an ontological approach to what makes us human, I propose in this chapter a *neurocognitive, interpersonal model of humanization*, which is deontological and non-essence based. In this model, humanization of another person is defined as a fundamental social brain ability, which is extremely precarious due to its "flexible" nature (Harris, 2017) and our tendency to spontaneously dehumanize various groups and individuals on an everyday basis. Humanness, dignity, human worth, or human uniqueness – however we decide to call it – is not an irreducible essence located in the individual like an invisible nugget, but a physically detectable, cognitive activity that takes place on an interpersonal level between people and groups who perceive each other through their social brains. Humanization is thus located in the complex and delicate interpersonal nexus between the perceiver and the one who is perceived. It is within this nexus where our political questions and theories about toleration, multiculturalism, and human rights need to be placed.

What this chapter does *not* set out to do is to argue against the moral and political aspirations that drive those philosophers, activists and politicians who fight for the recognition of universal dignity of all humans. Stating that dignity is neurocognitive and interpersonal, instead of deontological and non-essence based, does not mean that we cannot justify the defense of universal rights for leading a life unharmed by the state and fellow humans, or that we cannot argue for respect and humanization of all human beings. Rather, what the neurocognitive, interpersonal model of humanization introduces is an alternative, biologically grounded dimension of reality, where the phenomenon of

dignity is constructed around the neurocognitive capacities of the perceiver and the one who is being perceived.

Therefore, the second main argument that this chapter sets out to advance is that in order to resist and tackle political assaults on people's "dignity", we do not necessarily need an ontological proof of "dignity" itself. Rather, I argue that the existing evidence on the damaging effects of racism, sexism and dehumanization on targeted individuals' physiological health, executive brain functions and overall well-being and behavior is sufficient to justify a defense of "human dignity" (Dardenne et al., 2013; Gibbons et al., 2012; Harrell, Hall, & Taliaferro, 2003; Kteily, Bruneau, & Hodson, 2016). Therefore in order to oppose and fight political dignity violations, evidence showing how these violations can *harm targeted individuals in terms of their ability to fully function and thrive in their capacity as human beings* is both more persuasive and sustainable than ontological and essence-based proofs of dignity. To put it more bluntly, what happens between people – i.e. the interpersonal and neurocognitive ascriptions and denials of humanness to each other – is of greater political relevance than the existence of an irreducible human essence. Ultimately, what is at stake in political violations of dignity is not the eradication of dignity as an essence, but the threat that the frequent and often unconscious occurrence of dehumanized perception poses to a viable kind of liberal democratic politics.

4.2. Why Does Dignity Matter for Politics?

The idea of human dignity is both ubiquitous and elusive. This might be due to the fact that the term "dignity" has an air of immaterial fuzziness about it, or because

"dignity" has been used in almost every possible debate ranging from human rights (Amnesty International, 2005; Habermas, 2010; Lutz-Bachmann & Nascimento, 2014), animal rights (Cavalieri, 2001), euthanasia (Azize, 2007) and racial discrimination (Clark, 1970; Mitchell, 2005; Palmer & Smith, 2010), to the set-up of penal systems and the distribution of welfare (Chan, 2005). Within the Western post-Enlightenment discourse, the political question that lies at the heart of these debates can be summed up as follows: '*Does human dignity exist independently of a rights-guarantor (such as the state or international law), and if so, how can its existence be proven?*' When this question is posed, it is almost always assumed that dignity is a kind of invisible "essence" located in each human being (or possibly, also in non-human animals, see Röska-Hardy & Neumann-Held, 2012) and that therefore the ultimate goal is to demonstrate with a convincing proof that this "essence" exists. The problem of "dignity" is thus construed as an ontological one.

What drives the search for dignity's essence is the fear that without a non-negotiable, objective claim on the unique worth of being human, the future of universal rights and equality is under peril. For example, Jack Donnelly, a leading human rights theorist, states that we simply need to take a leap of faith when it comes to the existence of dignity (Donnelly, 2003). The fear is that without ontological claims laid on dignity, the demands for rights become mere subjective claims. The assumption therefore is that once an objective, ontology-based proof can replace a more instinctive moral belief about the inherent worth of human beings, the desirability and necessity of human rights can be justified cross-culturally and cross-nationally.

Indeed, from the UN Declaration of Human Rights (UDHR) to the *Grundgesetz* in Germany's 1949 Constitution, as well as the U.S. government's *National Security Strategy* spanning the last couple of decades, an essence-based idea of dignity continues to be invoked in both national and international politics. I argue that this can be traced back to the political question that I singled out at the beginning of this chapter, namely that in these above examples, governments and international institutions are trying to establish the existence of dignity independent to a formal rights guarantor. For example, the UDHR demands recognition of the "inherent dignity (…) of all members of the human family" and that "all human beings are born equal and free in dignity"[26]; whereas the German *Grundgesetz* (Basic Law), signed in 1949 in the wake of German atrocities committed during the war and through the Holocaust, states that human dignity is inalienable and untouchable by anyone[27], especially by state power. Similarly, the U.S. government's *National Security Strategy* (2015) lists "human dignity" numerous times together with the advancement of rights, prosperity and security; and further pits the preservation of dignity against the outbreak of conflict[28]. Within this list, dignity is the most intangible and perhaps least explicitly political value, yet one philosophical tenant underlying U.S. foreign policy seems to be based on the existence and importance of "human dignity".

Undoubtedly what these documents foremost try to invoke is the idea of natural inalienable rights, thus "dignity" serves more as a decorative rhetorical device for validating more substantial politico-legal ideas, rather than introducing a freestanding

[26] UDHR, Preamble and Article 1.
[27] Deutsches Grundgesetz, Artikel 1: "Die Würde des Menschen ist unantastbar."
[28] US National Security Strategy, February 2015, signed by President Obama.

argument about human dignity itself. However, in a more subtle fashion, these invocations are also making a distinct point about how dignity matters as an independent phenomenon, in that it is believed to be out there in the world and in need of protection, even if it might not have fully materialized into current political structures, language and rights[29]. What this chapter is asking is whether these philosophical invocations of an essence-based human dignity are in fact necessary to achieve the political goals that these governments, institutions and movements are trying to attain.

To take a step back, where might the understanding of dignity as in independently existent phenomenon have come from? In the Western context, Judeo-Christianity seem like an obvious starting point, consider *Genesis 1:26-27*: "Then God said 'Let us make man in our image, according to our likeness'…so God created man in his own image (…)". This is echoed in St. Thomas Aquinas' *Commentary on the Sentences* (1252), in which he states that "dignity signifies something's goodness on account of itself" (Aquinas, 1975), essentially affirming the intrinsic value of dignity, which it acquires through its fitting place in God's creation. In the Reformation period, John Calvin reasserts the link between *Imago Dei* and dignity by saying that "God's image is the perfect excellence of human nature" (Calvin, 1960, p.190). Indeed, "from the early Church to the sixteenth century to the present, appeals to the image of God and a concomitant assertion of human worth or dignity have been constants in Christian moral reflection about the person" (Howard, 2013, p.2f.).

[29] For a discussion of human rights *foundationalism* vs. human rights *pragmatism*, see Luban, 2015. The former contends that legal human rights should copy universal moral human rights; the latter believes instead that the semantics of "human rights" ought to correspond to its practical use and implementatory effects. In the context of this chapter's argument about dignity, a *foundationalist* argument most likely would support the idea that dignity is essence-based.

The dignity idea therefore has strong roots in Judeo-Christian thought – as a *substance* that exists by virtue of birth and possibly even precedes birth through its connection to the eternal *Imago Dei*, as a *principle* or human *status* that cannot be negotiated and should not be violated by worldly powers, and as a *value* that lives on unchanged despite political and constitutional transformations that are taking place now or in the future. Even when asking atheist or non-Christian human rights activists today, the account they give about the inalienability of human dignity (and the ensuing rights attached to it) often sounds strikingly similar to the above (Amnesty International, 2005; Blanco, 2014; Cruft, Liao, & Renzo, 2015; Novak, 1999). This is because in its core, the current human rights discourse, and the nations and international organizations that ascribe to it, embraced the philosophical tenets of the Judeo-Christian idea of dignity wholeheartedly, especially in the post-War context, in which most of the current legal vocabulary of human rights was developed (Glendon, 2001). If we go back to the question *'Does dignity exist independently of a rights-guarantor?'*, then an essence-based, Judeo-Christian idea of dignity allows us to answer the question with a resounding Yes, and the political and legal challenge that remains is to compel rights-guarantors to recognize and protect human dignity, and the rights that are attached to it.

That being said, the Judeo-Christian influence on current accounts of dignity is undoubtedly more complicated than suggested above. Immanuel Kant, whom Michael Rosen justifiably calls "that thinker on whose giant shoulders the modern theory of human rights largely rests" (Rosen, 2012, p.19) and whose "thought about dignity should stand at the center of any historical account of dignity" (Ibid.), famously tried to define human worth (*Würde*) not as defined through God or an external authority, but within an

individual person's capacity for autonomy and making one's own moral laws. Kant developed this idea in his *Groundwork to the Metaphysics of Morals* (1785), in which he states that dignity has an 'inner value' that cannot be replaced by a relative value, and therefore serves as an 'end in itself' (AK. 4:434-435). It is here where the foundation for Kant's *categorical imperative* is laid down, in which Kant demands that human beings should treat each other as ends in themselves. However, despite Kant's attempts at liberating the idea of dignity from its reliance on the external image and authority of a Christian God, Kant's idea of dignity as an intrinsic, non-negotiable value still retains some aspects of idea of an essence-based understanding of dignity.

The belief in an inviolable human worth, enshrined in natural rights, has been contested and even been subjected to ridicule in the history of political thought. Most famously perhaps, Jeremy Bentham's quote that "natural rights" is "simple nonsense: natural and imprescriptible rights, rhetorical nonsense, – nonsense upon stilts" (Bentham, 1848) comes to mind. Although Bentham is referring to natural rights here, I argue that the implicit reference here is to human dignity – Bentham contests the plausibility of the existence of human dignity, as implied in natural rights. Bentham continues saying that "and of these rights, whatever they are, there is not, it seems, any one of which any government can, upon any occasion whatever, *abrogate the smallest particle* [my emphasis]" (Ibid.). The reference here to the physical non-properties of natural rights, and human dignity as its foundation, are a strike against the ontological, essence-based claims underlying dignity. To Bentham, rights do not exist independently of rights-guarantors, whether that is a political-legal guarantor such as the state, or a moral guarantor such as society's majority.

One way to counter Bentham would be to point to his contemporary, David Hume, who impassionately argued in a much-overlooked chapter in the *Enquiry Concerning the Principles of Morals* (1751) that human beings have a natural tendency towards benevolence and empathic emotions towards their fellow human beings, and that this special 'sentiment' is the driving force of morality. Although Hume agrees with Bentham's skepticism of defining characteristics about human nature and virtues through abstract principles and metaphysics, he claims that in the empirical world, "morality is an active principle" (Hume, *ECPM* Gale Version, p.6), in which the driving force is a benevolent human sentiment that people in large parts have for each other, and which, most importantly, "nature has made universal to the whole species" (Ibid.).

In other words, humanizing others is an "activity", not a metaphysical essence, which exists for Hume in the world already and does not need to be proven. Hume rejects the cynical view that all human moral sentiments are driven by self-interest, but insists that he has observed in many instances where family and friends, and even strangers, have bestowed on each other benevolent feelings without ulterior motives. I will build on Hume's idea of moral sentiment as a cognitive-psychological activity instead of metaphysical essence, and his claim about the natural and universal occurrence of empathic feelings towards others in my neurocognitive model of humanization.

Human rights scholars and activists have pointed out that a Benthamian approach towards human dignity leaves individuals entirely at the mercy of governments and the tyranny of majority opinions, insisting that the affirmation of the existence of human dignity independently of a rights guarantor is essential to the purpose and success of human rights (McCrudden, 2008; Kateb, 2011; Kretzmer & Klein, 2002; Rosen, 2012;

Simmons, 2015; Tasioulas, 2012) – this position is commonly referred to as a "foundationalist" one within the human rights discourse. Critics of this discourse (Beitz, 2011) argue that international human rights do not need a philosophical foundation other than the practice of human rights, or that at the very least, the idea of human dignity can be decoupled from human rights applications, without making the latter less significant or successful (Buchanan, 2015; Luban, 2015; Waldron, 2015). This chapter sides with the latter camp, in that it maintains that what matters is not so much the independent existence of dignity but the acknowledgment that dignity violations at an interpersonal and neurocognitive level are harmful both to individuals and the viability of liberal democratic politics.

In a completely different vein to Bentham, Edmund Husserl, 20th century phenomenologist, introduced the idea of the *Lebenswelt* (lifeworld), in which intersubjectivity and empathy, as its foundation, play a crucial role. Discussed mainly in Husserl's *Cartesian Meditations* (1931) and *Husserliana* XIII-XV (1950), intersubjective experience is understood as empathic experience, which in turn makes up the 'lifeworld' of human beings. The 'lifeworld' is the pre-given and generally unreflected backdrop of our social world in which we make sense of ourselves and others – phenomenologically speaking, it represents the way members of social groups structure the world into objects. From a starting point of egocentric subjectivity, they move towards a psycho-physical perspective of the objective, spatio-temporal order (Beyer, 2015). In this 'lifeworld', my self-image as a full-fledged person comes about through what Edith Stein, Husserl's former Ph.D. student, calls 'iterated empathy' (Stein, 1917), meaning that I put myself into the position of someone else, under the condition that they will do the same, and that

the ability to ascribe intentional acts to someone else depends on a *bodily*, i.e. flesh-and-blood identification with each other as fellow human beings[30].

One cannot help but think of the strong parallels between the brain-based activity of mentalizing and humanization outlined in Chapter 3, and Husserl's phenomenological account of intersubjectivity and empathy. It is through Hume's observation of the natural and spontaneous occurrence of everyday humanization of others and Husserl's 'lifeworld' as intersubjective, bodily empathy that we have arrived at the entry point of the *neurocognitive, interpersonal model of humanization* that I wish to advance in this chapter.

4.3. A Neuropolitical, Interpersonal Model of Humanness

Why do we need to bring in the brain in the context of humanization? Firstly, by showing that the physical manifestations of (de)humanization at the brain level exist rebukes Bentham's challenge that claims to human dignity are groundless "non-sense" – in fact, what the brain data shows is that the perception of the humanness of others constitutes a fundamental way of how our social brains function on a daily basis, and that therefore the phenomenon of dignity is tangible, visceral and manifested in physical reality.

This reality, however, is not captured by brain imaging methodology as essence-based but instead as an interpersonal, brain-based experience. Furthermore, psychological and neuroscientific studies on how people react to stereotyping, rejection and ostracism cement the claim that dignity and humanization matters for those who wish to be

[30] See how Jewish phenomenologists in the latter half of the 20th century, such as Emmanuel Levinas, built on Husserl's idea of empathy post-Holocaust.

humanized (i.e. those at the receiving end), in that human beings have shown to be extremely sensitive to exclusion and denial of recognition by others (DeWall, 2013; Hiroshi, Taishi, & Mitsuhiro, 2015; Zadro, Williams, & Richardson, 2004). An alternative proof of human dignity could therefore be based on the fact that our social brains seem to have a deeply felt *need* to feel recognized, included and humanized, and that neurobiological and psychological evidence of this need might be sufficient for justifying political rejections of dignity violations.

The everyday and universal aspects of dehumanized perception can further allow us to rid dehumanization off its "evilness" and "sinfulness" connotation, by treating dehumanization not so much as a rare aberration from social and moral norms that is carried out by a small number of sadistic and evil people within the population, but rather, as an inherent cognitive ability that we all share and bring into politics. For example, the necessity for expanding the understanding of dehumanization as a more ubiquitous and widespread phenomenon came to the attention of Hannah Arendt, when she covered the Eichmann trail in Jerusalem (Arendt, 1963). Adolf Eichmann, a former German Nazi official who was tried in Jerusalem in 1961 for his implementation of the "Final Solution" strategy to exterminate Jews during the Holocaust, was portrayed by Arendt as an ordinary person who was not chiefly driven by sadism or a hatred of Jewish people, yet he was capable of and fully responsible for extraordinary crimes. What has by now become the famous title of her book, Arendt described the Eichmann phenomenon as the "banality of evil", capturing the everyday and ordinary aspect of how Eichmann executed and recounted his crimes against humanity. Arendt was baffled by the mundane and seemingly "normal" way how Eichmann came across, which foreshadows the

findings in the recent studies on everyday and subtle dehumanization by Leyens et al. (2001), as well as historical accounts of how Holocaust atrocities were committed by seemingly ordinary people (Browning, 1992).

Although Arendt's assessment of Eichmann was heavily criticized by various contemporaries (Abel, 1963; Ezorsky, 1966) and even discredited as having underestimated Eichmann's sinister intentions when he carried out the orders to exterminate millions of Jewish people (Robinson, 1965), her assessment of the "ordinariness" of Eichmann in her analysis of his *language* in the original German – in particular the impersonal jargon he used to describe key events as well as his lack of agency – illuminate how dehumanization can manifest itself in ordinary speech and accounts. What matters here is that *language usage*, as mentioned in Chapter 3, can serve as an indicator for the presence or absence of dehumanized perception. What struck Arendt most was Eichmann's inability to speak both about his victims *and himself* in a humanizing way when recounting past events, creating the image of the bureaucratic cog stuck in the workings of the larger machine.

To sum up, I propose the following model:

1. This chapter argues that in order to resist political dignity violations, we do not need an essence-based, ontological model of dignity. Instead, ***an interpersonal, neurocognitive model suffices in proving that human dignity exists***: both based on the human brain's ability to flexibly ascribe and deny humanness to others, as well as on the cognitive need to feel humanized, included and recognized by others. What matters is ***how dignity is negotiated, denied and demanded at the brain level between people***, not whether it exists independently of governments and laws as an irreducible essence.

2. The ontological hunt for dignity is misguided, because we *already* spontaneously humanize and dehumanize people on a daily basis. Trying to prove the existence of an inherent human essence is misguided on two grounds: firstly, it denies the fact that spontaneous humanization (in the form of spontaneous mentalizing, or

through feeling empathic emotions) of others is an inevitable and a basic part of how our social brains function,

3. And secondly, it neglects the mechanism underlying everyday, spontaneous *de*humanization, in so far as an ontological 'essence-based' theory of human dignity underestimates the way how everyday dehumanization permeates all aspects of everyone's social lives, making us all regular 'violators' of each other's humanness (even the most morally perfected person can still succumb to dehumanized perception), and that one cannot expect that once the existence of the 'dignity essence' is proven, dehumanization does not take place anymore.

4. But also, it underestimates the sheer **frequency and pervasiveness** of when a denial of individual human worth occurs and what affect this might have on those vulnerable groups whose dignity is affected by everyday dehumanization. Here, the relatively new field of everyday dehumanized perception is supported by more established research on automatic stereotyping and exclusion, see compelling data on frequency and pervasiveness of implicit biases gathered through the "Implicit Association Test" (IAT)[31].

5. Cultural neuroscience on **biculturalism**: difference in sense of self/self-construal amongst East Asians (interdependent) and Westerners (independent) suggests that the Western individualist-construal of the self is not the only path for people to define their human worth. But also, see non-Asian minority groups such as Muslim women in Western Europe: for example, an interdependent self is implied in French Muslim-headscarf debate, i.e. anti-headscarf feminists and philosophers such as Luc Ferry want to 'save' Muslim women from this 'undignified' interdependent construal of their self, by forcing a public space where they have to publicly display what is seen as a dignified, i.e. 'independent' assertion of the female body (Winter, 2008). In addition, other non-Christian religions such as Buddhism have a completely different idea of human dignity, stressing non-individualism as highest standard of the Self (Wu et al., 2010).

6. Bicultural individuals, if primed, can effortlessly switch between cultures and different self-construals and value systems (Chiao et al., 2010). How does this affect a monolithic, single-ontology based proof of dignity (even if it were found one day), if people can exist and switch between different ontological accounts of self? How can universal theories of human dignity be defended against the pluralistic experience of dignity in multicultural individuals?

---

[31] See AG Greenwald, DE McGhee, & JLK Schwartz (1998) "Measuring individual differences in implicit cognition: The implicit association test" *Journal of Personal and Social Psychology 74*: 1464-1480. The IAT has since been tested in a multitude of international settings, and has been strengthened through correlating fMRI data, see Kristine M. Knuston, Linda Mah, Charlotte F. Manly, & Jordan Grafman (2007) "Neural correlates of automatic beliefs about gender and race" *Human Brain Mapping 28*: 915-930.

Ascribing a human mind to others and trying to figure out what others think and feel as unique human beings within our immediate environment is a cognitive activity that we develop in infancy and continue to employ in a wide variety of social situations throughout our lives. This cognitive activity is called mentalizing or Theory of Mind amongst developmental psychologists and social neuroscientists (Adolphs, 2006; Wagner, 2015; Wimmer & Perner, 1983). Making sense of other people's minds, i.e. making sense of others as fellow human beings is of such importance to us that our brains will try to use various cues ranging from someone's facial expression (Adolphs et al., 1994; Calder et al., 1996), emotions in their eyes (Baron-Cohen et al., 2001), body language, to actively probing others and analyzing their self-ascriptions. Attributing mental states to others has been reported to develop in infancy cross-culturally (Avis & Harris, 1991), i.e. it is a universally occurring ability and need of the human social brain. In this sense, Hume was correct to insist that humans direct pro-social cognitive and emotional resources towards each other on a daily basis, and that this is a fundamental way in which humans operate within society.

Furthermore, the most recent research on mindreading draws strong links between mindreading and empathic reactions (McCall & Singer, 2013), as well as the moral evaluation of one's own and other people's actions, with some going so far as to claim that the "primary service of mental state reasoning may be for moral cognition and behavior" (Young & Waytz, 2013). Again, Hume's assertion of morality as an activity – not an ontological essence – grounded in the benevolent sentiments towards others resonates strongly with the neurobiological manifestations of mentalizing, pro-social emotions, and empathy.

It is important not to simplify the phenomenon here: the fact that I am ascribing mental states to others on a frequent daily basis, and that this might primarily serve moral reasoning, as well as result in Humeian 'benevolent sentiments' does *not* have to automatically result in a world where people live in harmonious unison as each other's friend. The first thing to keep in mind here is to separate cognitive activities, pro-social emotions and attitudes (i.e. what goes on in the social brain) with actual actions and behavior towards people (i.e. what goes on in the social world), which was already discussed in the discussion on neuropolitical methodology and epistemology in Chapter 2. In addition, cultural neuroscience stresses that in the context of bicultural individuals that *attitudes do not equal behaviors* (Boski 2008; Ward & Kus, 2010), in the sense that the way people navigate the cultural divide within their bicultural identities can be 'multideterminate'. On the question of multideterminate outcomes of psychological attitudes and neurocognitive abilities consider political scientist Robert Jervis, who repeatedly pointed out the self-deceptions and contingences at play in political perceptions and decision-making:

> "We cannot understand systems by summing up the characteristics of the parts or the bilateral relations between pairs of them (…) intuitively, we often expect linear relationships. If a little foreign aid slightly increases economic growth, then more aid should produce greater growth. But in a system a variable may operate through a nonlinear function. That is, it may have a disproportionate impact at one end of its range (…) Similarly, the effect of one variable or characteristic can depend on which others are present."
>
> (Jervis, 1997, p.34-35)

Jervis argues that different variables do not simply "add up" to one single outcome, but that it is the interaction between the variables (and other external, often unforeseen factors) that lead to a multideterminate outcome.

Whereas Jervis is describing what he calls "system effects" for situations on a national or international political scale, this chapter is concerned with a more micro-based framework of interpersonal ascription and denial of humanness and human dignity, from the perspective of our social brains. Jervis' system analysis can however be applied to our case here, in that dignity is a complex system in itself, which manifests itself at levels of neurobiological, attitudinal, affective, as well as social behavior and political action. Treating 'dignity' in the one-dimensional, essence-based way (which also leaves out behavioral outcomes) as is currently the case within many debates in Western moral philosophy and theories of human rights, is insufficient in grasping the "system complexity" at stake with humanizing others, and eventually the challenges of putting humanizing attitudes into actual actions and practice.

4.4. Dehumanization's "System Complexity"

4.4.1. Dehumanization's "System Complexity" I: Perceptions, Attitudes, Behavior

Research into subtle, everyday dehumanization (Leyens et al., 2001) has shown that our social brains are highly susceptible to quickly dehumanizing others as a result of disgust reactions, skin color and out-group membership, amongst various others triggers. The previous chapter already introduced the concept of dehumanized perception at the brain level, trying to explain why and how it matters or the politics of dealing with out-groups. To recap, dehumanized perception is chiefly characterized by an absence of mentalizing (i.e. inactivity in the *medial prefrontal cortex* and *temporoparietal junction*), and that it can manifest itself in animalistic or mechanistic dehumanization of others (Haslam, 2006). Dehumanizing others is thus a way of how we navigate ourselves

through our social world, where evolutionarily we had to make quick decisions about who belonged into our in-group (Brüne & Brüne-Cohrs, 2006). Behaviorally – if we do end up manifesting behaviors based on spontaneous, rapid dehumanization of others – this can result in both active and passive harm to others: as a result of dehumanized perception, we can end up demanding harsher punishments of offenders in the penal system (Vasiljevic & Viki, 2014), poorer health care quality and insufficient administration of pain medication towards Blacks (Hoffman, Trawalter, Axt, & Oliver, 2016), to the more obvious behaviors associated with dehumanization, such as extreme acts of aggression and atrocities committed during civil wars and genocide.

How does this relate to dignity? One premise on which the neurocognitive, interpersonal model of humanization rests is that dehumanization is known to occur spontaneously and with high frequency in our lives. Not all neurocognitive experiences of dehumanized perception result in active behavior, even passive behavior can only arise if the socio-political circumstances encourage neglect towards a dehumanized out-group: for example, if I were to dehumanize homeless people, my neglect behavior towards them might only arise and become significant if I take part in decision-processes that allocate welfare resources to various groups in society, and as a result of dehumanization of the homeless, I forget to take their needs into account when I make those decisions. Again, as with humanization, dehumanization has to be considered in the context of its own "system complexity".

What I wish to highlight in this chapter, however, is that the automatic and frequent nature of dehumanized perception gravely undermines an essence-based, ontological proof of dignity's existence. People cannot simply 'decide' that they

recognize the existence of human dignity and as a result, never deny humanness to someone again, but rather, no matter how much we pledge to respect someone's dignity, chances are high that we might deny unique humanness to them if dehumanization is triggered by disgust, racial or other stress factors. Essence-based accounts of dignity therefore need to take into account the neurocognitive capacities of social actors for both humanization *and* dehumanization, and develop a more nuanced account of how 'dignity' is negotiated in the social world between social brains.

Finally, we again need to take into account the perception vs. behavior divide. One could say that human rights activist are well-aware of people's capacity for dehumanizing others – criticizing those who violate humanness is a human rights activist's core objective – yet what they chiefly focus on is a *behavioral* manifestation of dehumanized perception, i.e. atrocious actions and discriminating decisions. Dehumanized perception, however, as explained above, is not just defined by behavior but originates first in a neurocognitive perception, is then felt as an excluding attitude and accompanied by emotions of disgust, stress or fear, and might only as a last step materialize finally in observable behavior. The model put forward in this chapter suggests that political scientists and human rights theorists take into account not just behavioral manifestations of dehumanization, but also consider the perceptual, attitudinal and affective dimensions, and think about how those do (or do not) interact with specific behaviors in the social and political world.

4.4.2. Dehumanization's "System Complexity" II: Culture and Hyperdiversity

Moving onto cultural neuroscience, this model also tries to incorporate the identity experiences of bicultural individuals and non-Western people within Western liberal democratic societies. Today, an ever increasing number of people hold multiple identities within them, which contributes to increasingly hyperdiverse societies in Western liberal democracies (European Commission, 2013; U.S. Census Bureau, 2010). In the context of universal claims to human dignity, a monocultural and exclusively Western understanding of the Self can become problematic for the following reasons. For example, in studies on differences in self construals amongst Westerners and East Asians, cultural psychologists have proposed for a long while that Westerners have an 'independent' idea of self, whereas East Asians have an 'interdependent' idea of self, which is much more tied to one's community and family (Markus & Kitayama, 1991). In the last decade, cultural neuroscience has managed to provide supporting evidence for this claim through fMRI data that shows that the brain area that is usually implicated in self-reference is activated in both Westerners and East Asians when prompted to think about themselves, but when both are asked to think about their mothers, only in East Asians does that same brain area for self-reference activate as well – meaning that an East Asian person views their mother as identical to themselves, which confirms the psychological theory that East Asians have an "interdependent" sense of self (Zhu, Zhang, Fan, & Han, 2007).

Similarly, in the realm of religious identity, a comparison between Christians and Tibetan Buddhists showed differences in their self-construal, detectable at the brain level. When asked to think about themselves, Christians activate their *dorsal medial prefrontal cortex*, which is usually active during third person perception – this could be explained

by the fact that Christians routinely judge themselves through the eyes of God (who looks onto them as a third person) (Han et al., 2008). Buddhists, on the other hand, do not activate any brain region usually implicated in self-reference *at all*, which could be explained by the fact that the daily religious routine and the core of its religious philosophy tries to rid the Self from an ego-based, self-referential concept, and to this effects employs various meditation and prayer techniques to exculpate the Self from self-reference (Wu, Wang, He, Mao, & Zhang, 2010).

It is remarkable that the differences of self-construal in different religions can manifest themselves in brain activity and possibly even brain structure. This poses a serious challenge to *monocultural constructions of human dignity*, as they are currently employed within much of contemporary Western philosophy (Taylor, 1989). The picture becomes even more complicated when considering individuals who hold multiple identities within themselves, i.e. who are both Western *and* East Asian or who identify with a secularized Judeo-Christian individualist conception of rights but are at same time devout Buddhists. In the burgeoning field of bicultural brain science, researcher have found that Asian-Americans, for example, can switch effortlessly between their independent Western and interdependent East Asian identity, if primed with a specific cultural symbol (Asian dragon; US flag), a culturally-specific story or language. In other words, two completely different neurological circuits activate depending on how bicultural individuals are primed (Oyserman & Lee, 2008). In another experiment more directly relevant to political science models, the level of cooperation in Asian-Americans increased in the Prisoner's Dilemma, when primed according to their interdependent Asian identity (Wong & Wong, 2005; Utz, 2004).

It is important to note here that many of these insights could not have been made without fMRI and brain science tools, since survey and attitudinal methods (which are usually used in psychology and social science research) in the form of self-reporting might not always be accurate. For example, Warnick and Landis (2015) stress "it is important that intercultural behaviors are measured directly and not through self-report questionnaires, which are necessarily affected by memory and demand characteristics" (Warnick & Landis, 2015, p.17f).

I propose a neurocognitive, interpersonal model of humanization vs. an ontological, essence-based model of human dignity because it manages to incorporate the ongoing demographic shifts in Western liberal societies, where an increasing number of people hold bi-or multicultural identities and define their sense of dignity through more than just a single, Western monocultural framework of what it means to be human. It is this effortless switch between multicultural identities that Western monocultural models of dignity within current theories of toleration and human rights often fail to capture. In addition, in the context of cosmopolitanism theory (Brown & Held, 2010), my model suggests that cosmopolitanism is not just a "value" out there in the world, or a lifestyle decision, but that many people in our increasingly hyperdiverse societies embody cosmopolitanism at the brain level, in oftentimes conflicting and confusing ways. Identity hybridization can create multideterminancy of behavior, which underlines the need for current conceptualizations of "dignity" to take into account the cultural "system complexities" that come with it.

4.5. Multiculturalism and Toleration

So far this chapter made a strong case for the Judeo-Christian origins of "dignity" but equally, one could go back to Antiquity, where "dignity" signified one's designated status in the world. In Cicero's *De Officiis* (On Duties), he linked one's *dignitas* to the "honored place" which one was holding in the social order of things, but also, perhaps more radically, Cicero proposed in addition to this that human beings possessed *dignitas* because of human being's superiority over other animals (Cicero, 2014). Humans, in other words, were thought to hold a special status place within the world (and universe – within the framework of Cicero's Roman cosmology), i.e. there was something distinctive about their existence that merited an elevated position from other living beings around them.

The idea of the 'distinctiveness of man' carried through the history of Western philosophy up until the Renaissance, where Pico della Mirandola reiterated this point in *De Dignitate Hominis*. Mirandola claimed that humans were special because they did not just fulfill a preordained role, but were given the capacity by God to choose their own destiny (Mirandola, 1953). This leads Michael Rosen to argue that Mirandola "opens a relatively clear path toward the use of dignity we find in modern human rights documents", turning it from "a matter of the elevated status of a few persons in particular society to being a feature of human beings in general" (Rosen, 2012, p.15).

The link between status and dignity was redrawn for modern political thought by Jeremy Waldron (2012), who argues that dignity in modern times constitutes a set of rights based on an underlying idea of the status in question, and that dignity ought to be defined as a status-concept, not a value-concept. Likewise, Mirandola's compelling theory about status-based dignity grounded in human autonomy has found its way into

other modern liberal theories, such as multiculturalism. In the multiculturalist view of society, various cultural, ethnic and religious groups all can lay a claim to the inherent dignity (and therefore a claim to be tolerated or even protected by a rights guarantor, such as the state) of their group identity. Dignity here is both pluralistic and unitary in nature, in that all groups can potentially lay claim to their inherent dignity, but the status (i.e. the particular place in society) of each group in relation to the majority group can vary greatly.

One of the most influential works in this area is Will Kymlicka's *Multicultural Citizenship: A Liberal Theory of Minority Rights* (1995), in which he justifies the recognition of minority group's inherent dignity (and subsequent rights attached to it) through the idea of the good life, meaning that in order for people to achieve conditions of the good life, they might want to draw on the values, customs and beliefs of their group identity. The chief criterion through which the good life can be categorized is through autonomy: people should be allowed to lead their lives from the inside, i.e. in accordance with what they believe gives meaning and worth to their lives, but most importantly, they should always be allowed to question and opt of their chosen beliefs, i.e. be given the autonomy to choose what ultimately is the good life for them (Kymlicka 1995, p. 81).

As others have argued (Parekh, 2000), Kymlicka's definition of autonomy is less ambitious than those of John Stuart Mill, Kant or Joseph Raz, in that it does not aim for a substantial kind of self-fulfillment in terms of self-legislation or self-creation. Autonomy for Kymlicka is simply a necessary condition for the good life to thrive, which is why he judges minority groups not according to the actual contents of their identity, but whether

they allow their members to develop into autonomous individuals, who can, if necessary, challenge beliefs held within their group. The deciding factor for Kymlicka is not whether they can opt out of their identities, but to be given the freedom to move within it.

Kymlicka's position aligns with liberals such as Michael Walzer, who point out the pervasiveness of "involuntary association", i.e. familial, social, birth and other group memberships, which we are all part of throughout our lives and are often born into without initial choice (Walzer, 2005, p.2). Contradicting Rousseau, Walzer contends that "we are not born free, we are not born equal" (Ibid.) – echoing the multiculturalist view that group recognition depends on a group's status. In contrast, other more purist liberals such as Joseph Raz justify culture on more functional grounds, demanding that one's affiliation and love for one's cultural identity has to be "rational and valid" and be based on the "right reasons" (Raz, 1994, p.184). Raz further claims that a secure grounding within one single cultural group "is a precondition" for individual freedom (Ibid., p.178). In a much more radical and polemic vein, Jason D. Hill, a self-proclaimed radical and unapologetic liberal, demands that if a minority member does not adopt individualist values in the strict liberal sense as defined by Hill, then they affirm the "uncivilizedness" of their identity-belonging and therefore automatically loose the right to be accorded respect and entry into the liberal political community (Hill, 2009).

The neurocognitive, interpersonal model of humanization and its insights into the actual cognitive mechanisms of self-construal and identity-switching within bicultural individuals offers us powerful evidence against Kymlicka, Walzer, Raz and Hill. First of all, liberal multiculturalists' preoccupation with autonomy as the guiding principle for judging the worth of a cultural group is entirely based on a (ironically) monocultural

Western model of individualist identity. They do not take into account that interdependent identities such as Confucianist and East Asian identities can derive equally as much dignity as independent ones, and deserve to be respected for what they are. At the brain level, someone who constructs their sense of self – and arising from that, the existential, moral and emotional meaning of their lives and the relationships they have with others – in an interdependent way activates distinctively different neuronal circuits to someone with an individualist sense of self. This interdependent representation at the brain level is thus not simply an erroneous or misguided choice (as Raz would have it) but a physical-biologically experienced reality that is deep-seated and fundamental.

That being said, the insights into how bicultural Asian-Americans can switch effortlessly between independent and interdependent identities, if primed, suggests that as much as identity affiliations are neurologically represented at the brain level in a biologically significant way, the brain also has the flexibility to incorporate new identities and activate these effortlessly if primed. It also shows that in the most influential theoretical accounts within liberal multiculturalism, the depiction of the issues at hand are way too simplified or in some cases, outright wrong (curiously, many of the leading multiculturalists are monolingual, or do not fluently speak a language of the Global South or the minority groups of color they theorize about). It certainly proves wrong Raz's claim that one can only have one identity group in which one is grounded – indeed, in the light of the dramatically shifting demographics and the increasing number of bi-and multicultural individuals within societies around the world, Raz's claim does not just seem outdated but deeply dehumanizing.

In recent years, Raz's positions have been criticized by a more subtle take on multiculturalism and minority experiences, such as in the edited anthology *Minorities within Minorities* (2005) by Avigail Eisenberg and Jeff Spinner-Halev. In there, Rob Reich, citing Jeremy Waldron on cosmopolitan individuals with highly diverse biographies, agrees that "the phenomenon of multicultural, hybrid individuals thus challenges Raz's liberal multiculturalism" in so far as clear cultural boundaries become blurred, people are not Herderian wholes, and "people construct their identity and exercise their autonomy within a multiplicity of cultural frameworks" (Reich, 2005, p.222). The way Reich counters Raz through examples of individual biographies of hybridized people is persuasive and evocative, yet additional brain evidence can help build a more complete picture of the cognitive mechanisms at play for bicultural individuals. For what the biographical stories do not tell us is how and when exactly bicultural individuals switch their identities – or indeed, if they switch them at all. The brain science is important here not because it provides us with an ontological definition of who hybridized individuals are, but how certain crucial cognitive mechanisms function within them.

I now wish to apply the neurocognitive, interpersonal model of humanization in the context of toleration. Michael Walzer feels strongly about minority groups who do not pull themselves up by their bootstraps and fail to overcome, in Walzer's own words, the "permanent state of suspicion", in which this "nervous", "obsessive", "self-defeating" and angry minority group claims that it is not granted respect (Walzer, 2005, p.37), instead of trying to gain respect by becoming economically successful (I claim that Walzer is referring here, in a thinly veiled way, to low-income African-Americans) in

society[32]. For Walzer, economic status is intrinsically linked to dignity ascription, and those groups who lack respect in society have simply not sufficiently understood this fundamental connection. Walzer calls this "meat-and-potatoes-multiculturalism", where the "material strengths of groups compels their mutual respect" (Ibid., p.38). It is puzzling where exactly Walzer derives his economic status optimism from, given various historical examples in recent history, such as the fact that neither high economic status nor educational achievements of European Jews in the 19th and 20th century protected them from subsequent persecution in Europe (Albanis, 2002; Brustein, 2003). It is exactly this paradox of wealthy, middle-upper class Jews who were respected in their town one day, and harassed and deported by their neighbors the next that Walzer seems to be oblivious about.

But even in less extreme political contexts, the social psychology and neuroscience of dehumanization shows that in the U.S., Jews and Asians are routinely mechanistically dehumanized and stereotyped with emotions of envy, in that they are seen as economically competent but humanly lacking in warmth (Fiske, Cuddy, Glick, & Xu, 2002; Haslam, 2006). Walzer omits the long tradition on stereotyping research within social psychology of the last three decades, such as for example Fiske et al.'s "stereotype content model", which shows that stereotypes are often mixed (i.e. a group can be seen as competent in one way, but lacking in humanness and warmth on the other hand – or the other way around). An important take-home point about the envy stereotype category is that the behavior resulting from it is aggression (Cuddy, Fiske, & Glick, 2007).

[32] The 'pull yourselves up by your bootstraps' argument was used already commonly employed in the US in the 1960s, in the context of shaming and blaming the urban poor, see sociological study by Gans, 1995.

If interpreted less generously, we could treat Walzer's materialistic dignity argument as a manifestation of his own mixed stereotyping towards certain minority groups, if interpreted more generously, we could say that his economic status reasoning is deeply misguided. In the case of the latter, Anna Galeotti, a recent critic of liberal theories of toleration, would counter that the problem here is that "liberal theory understands toleration [as] a conflict produced by differences which usually pertain to individual choice" (Galeotti, 2002, p.5). Leaning on Amartya Sen and Martha Nussbaum's capabilities approach, she argues that minorities do not just possess fewer resources and opportunities, but have fewer capabilities altogether. In other words, "those whose collective identity is despised, or who are subject to prejudice and stereotyping, usually experience a lack of confidence or self-esteem and various forms of self-hatred, which make it much harder for them to become fully functioning social agents and citizens" (Galeotti, 2002, p.9).

For Galeotti, liberal theories of toleration fail to grasp unequal status of social groups to start with, which is not simply (just) economic inequality as Walzer would have it, but a "special kind of power which the majority can wield over social standards, standards so deeply rooted that they have been embodied in political institutions without our even being aware of it" (Ibid., p.227). In the neurocognitive model of humanization, this "special power" can have particularly wide-reaching effects if it is accompanied by dehumanized perception, i.e. absence of mentalizing and mindreading, of the minority group. As I argued in Chapter 3, neurocognitive capacities have to be contextualized in political and economic systems of inequality, in order to grasp the full extent to which

dehumanization (and to whom exactly) can be damaging, and in order to determine whether subtle, everyday dehumanization is of concern in a particular situation.

Galeotti stresses the largely unconscious ways in which majority standards have skewed a sense of equal recognition for minorities within political institutions, and the additional knowledge we have about the occurrence of everyday dehumanization, which is why national and transnational political institutions should pay special attention to the disadvantage minority groups face as an effect of this. The European Convention on Human Rights[33], however, does *not* have special minority rights provision, i.e. there exists no requirement for positive rights – recognition of minorities' demands and needs is usually based on the argument of "democratic pluralism" (which implies a certain basic equality between all groups, which does not exist *de facto*) (Pentassuglia, 2012; Peroni, 2015). How can the European Convention make special provisions that would facilitate the humanization of minorities and elevate them to equal cognitive status as others? The European Convention, and other transnational institutions, need to recognize that the humanization of minority groups faces particular humanization challenges, which should be taken into account in potential future minority provisions.

4.6. Conclusion: Becoming Autonomous Agents of our Social Brains

In this chapter I tried to advance a concept of human dignity that is not based on a transcendental, inherent essence but instead on the neurocognitive processes and interpersonal dynamics at play in societies that are becoming increasingly diverse. This leaves one to wonder whether the identity landscape of the 21st century has become too

[33] http://www.echr.coe.int/Documents/Convention_ENG.pdf

complex to navigate oneself through it morally and politically. Should we just try to avoid dehumanized perception, even though this might be impossible and even if latest research suggests that too much conscious humanization of a vulnerable out-group might lead to increased dehumanization instead, because we would rather not feel someone else's pain? How is one to conduct one's political brain, given the insights gained in this chapter?

In conclusion, I propose – leaning loosely on Kant here[34] – that the most important duties we owe in terms of our neurocognitive perceptions are not towards others, but ourselves. This is to say that we have to become *autonomous agents of our own social brains*, before we can engage in purposeful humanization of others. Cameron, Harris & Payne (2016) showed that anticipated emotional exhaustion leads people to dehumanize groups in order to avoid emotional pain. Therefore instead of asking people to actively humanize everyone, which this study suggests might overwhelm the emotional capacities of our brains, practicing "mindfulness" whilst we dehumanize an out-group might help in lessening dehumanizing reactions the next time we meet them. Although this does not equal overcoming dehumanization completely, it could potentially blunt aggressive behavioral outcomes (Borders, Earleywine, & Jajodia, 2010; Heppner et al., 2008). Becoming autonomous agents of our social brains is no small feat. It requires that people view their moral and political obligation to others as a neurocognitive understanding and mastery of their own neuronal humanization and dehumanization networks. If we can accept that our brains are major players in the ascription of someone

[34] Kant was no humanist in that he believed we owe our most important duties to ourselves, see AK 4:435 and AK 5:86, also Rosen, 2012, p.138ff.

else's dignity and humanity, we can begin to putting our brain abilities at the center stage of interpersonal acts that involve ascription and affirmation of someone else's "dignity".

5

5. Dangerous Dehumanizing Divides: Neuropolitical Effects of the "Civilizationary Clash" Theory

> If mankind is to get rid of the coercion to which the form of identification really subjects it, it must attain identity with its concept at the same time.
>
> Theodor W. Adorno, *Negative Dialectics*

5.1. Alive and Potent: The 'Civilizationary Clash' Theory

In this chapter, I apply the PBP on dehumanization in the realm of International Relations (IR), particularly to discourses of "clashing civilizations" and "civilizationary divides". This chapter examines the theories behind these discourses, specifically in regard to how they operate on either subtle or blatant dehumanizing notions of "uncivilized" or "barbaric" out-groups. I argue that locating the source of political conflict and challenges to governance-building along *civilized vs. barbaric* lines is problematic on two accounts: firstly, due to how this particular kind of distinction can lead to the reinforcing of dehumanized perception of other out-groups within our brains, and secondly, how describing certain groups as barbaric or lacking civilized capacities can lead to feelings of hostility and increased likelihood for intergroup violence.

Whereas the previous two chapters primarily focused on the political consequences of dehumanized perception in the brain of the excluder, this chapter switches perspectives by paying special attention to the implications of feeling dehumanized *by* someone else. In particular, this chapter discusses the concept of *metadehumanization* and how groups who believe that they are viewed and treated as barbaric will more likely resort to retaliatory violence (Kteily, Bruneau, & Hodson, 2016) and choose retributive over restorative justice options (Leidner, Castano, & Ginges, 2013). It is based on this that I argue that theories and rhetoric employing the concept of incommensurable civilizationary divisions are counterproductive for promoting

intergroup cooperation and peace, and might in fact exacerbate conflict and violence. This chapter urges political theorists and international relations scholars to reconsider the language and frameworks within which they frame international political conflicts, arguing that the employment of hierarchical and divisive civilizationary concepts is problematic because of its potentially dehumanizing priming effects on our social cognition, and of how it increases the likelihood of violent intergroup conflict.

Political scientists therefore need to be aware of how civilizationary rhetoric can negatively affect our cognitive abilities in viewing other groups as fully human, as well as the dramatic effects it has on intergroup attitudes and behavior. Instead of viewing the civilizationary argument as an insignificant rhetorical flourish in political exchanges and theories about the international world order, it should be treated as a central and decisive element in determining the outcome of peace, and intergroup cooperation and solidarity in international conflicts. In addition, the ongoing success and popularity of the civilizationary clash theory, both within academia and policy circles, but also amongst surging extreme right-wing movements across the Western world, could be ascribed to how this particular perspective towards out-groups feeds into and amplifies our natural tendencies towards in-group favoritism (Cikara et al., 2014; Krill & Platek, 2009).

The two most notorious works on the future of the global order that emerged since the end of the Cold War are most arguably Francis Fukuyama's *The End of History and the Last Man* (1992) and Samuel P. Huntington's *The Clash of Civilizations and the Remaking of World Order* (1996). In both of these works, the authors rely heavily on the concept of 'Western civilization', and how it relates to (and oftentimes contradicts) other 'non-Western' civilizations. Fukuyama lays the theoretical groundwork for Huntington's

thesis by arguing that world history's ideological struggles have come to an end with political liberalism as the only viable and persuasive ideology left. Building on this line of reasoning, Huntington posits that if the dialectical battle between political ideologies has come to an end, then what remains as the major friction point between countries and societies are their 'civilizational' differences. In particular, Huntington warns of trying to impose 'Western' values and ideals upon 'non-Western' societies, since for him 'civilizations' are bounded and often mutually exclusive due to fundamental and essence-like differences.

His controversial theory has been enormously influential in the post-Cold War debates on modernization, development, and within US foreign policy in particular, but also garnered a wide array of critical responses from postcolonial historians, sociologists and political scientists. For example, IR critics have argued that societies are never actually as bounded, insular and static as portrayed by Huntington – both in their histories and the way they transmit and communicate their cultural values and practices to the outside world (Hall and Jackson, 2007; Hobson, 2004); building on the earlier work of 20th century sociologists such as Pitrim Sorokin, Norbert Elias, Shmuel Eisenstadt and Benjamin Nelson, who all pointed out the importance of civilizational encounters, the porousness of borders and frontiers, and civilizations as dynamic, not static, entities[35]. Others have made a powerful case that 'civilizations' cannot be understood as actual actors on the global stage (Mazlish, 2001), since they are rarely unified entities in the

[35] For a review of the civilizationary discourse in the social sciences after Huntington's publication see O'Hagan (2007). For a recent re-assessment of the influence of Huntington's theory and the crucial role of the media in its initial promotion see Bantimaroudis (2015).

same way as states or cultures, nor do they hold the monopoly on political and military power, as global political actors usually do (Melleuish, 2000).

In the specific case of "Western" versus "Islamic civilization", others have joined this line of reasoning by pointing out that the United States' close relationship with Saudia Arabia and other Muslim states, as well as the amount of time that many Islamic extremists have spent in the West attests to the ambiguity of this alleged civilizationary divide (Berman, 2003), suggesting a more dynamic interdependency rather than an actual cultural clash (Said 2001). From a materialist viewpoint, postcolonialist theorists such as Tariq Ali (2002) argue that it is chiefly economic inequalities, and the structures that sustain them, which drive the decisions of political actors on both sides of cultural divides.

Remaining in the context of Islam versus the West, Fareed Zakaria (2001) believes that the frustration and anger expressed in Islamic terrorism does not originate within Islamic religion itself but stems from disillusionment with the West, as the Arab world has largely failed to implement in-depth modernization for its populations[36]. Continuing in this vein, Daniel Chirot (2001) claims in direct response to Huntington that the frictions in contemporary politics are less due to an incommensurable clash of cultural and religious values between societies, but stem from a materialist difference in levels of modernization between countries. It is this difference of development, and the ensuing variation in socio-economic and political progress, which causes clashing interests and beliefs between states. Amartya Sen (1999) echoes this sentiment in his own

[36] Indeed, the Arab Spring revolutions, which began in 2010 and gripped countries from Tunisia to Egypt, serve as example that neither Muslim nor Arabic societies are inherently incapable of democratic activism or the pursuit of liberal individual rights.

response to Huntington, when stating that "the practice of democracy that has won out in the modern West is largely a result of a consensus that has emerged since the Enlightenment and the Industrial Revolution" (16), thus stressing the contingent circumstances that aided the West in achieving stability and prosperity.

What Sen is trying to point out is that the successful implementing in the West of philosophical and revolutionary ideals in the real socio-political world was not due to an unwavering commitment to these very ideals over millennia, but a fortuitous culmination of historical circumstances and developments happening at the right time, in the right place, within a brief time period in very recent history. Therefore, non-Western societies that have not reached the same level of current Western standards of rights, governance and development are not necessarily in that position because of a lack of appreciation of ideals about the individual, equality and diversity, but rather, might be experiencing the crippling after-effects of colonialism and their own particular historical predicaments, in which political and economic progress did not coincide in the same fortuitous way as in the past three centuries in the West.

Huntington's view is commonly contrasted as pessimistic, versus Fukuyama's more optimistic assessment of the viability and endurance of liberal democracy (Mueller, 2014). However, they do converge in their shared belief in a *hierarchical* ordering of different global societies, in which those who have fully reached economic and political modernization (i.e. mainly Western societies) are implicitly seen as superior to others[37].

[37] In defense of Huntington, he did not explicitly state in his work *The Clash of Civilizations* that Western 'civilization' is superior to others. However, it can be argued that the likelihood of Huntington being completely unaware of which kind of historical and philosophical discourses he was inserting his civilizationary clash theory into (and the potential polemical and superiority-based conclusions others would draw from it) is very small.

In fairness to Fukuyama, it ought to be pointed out that unlike Huntington, he does not ascribe the divergence of development to innate and insurmountable cultural differences so much as to institutional variation. In his recent publications (2011, 2014) on the historical comparison of the developmental difference of Western and non-Western countries (and variation within Western countries' own histories), he argues that divergence in institutional development is the key variable for explaining civilizationary difference. The effort to establish impersonal, rule-based political order – in the Weberian definition of modern institutions and state power – is described by Fukuyama as an arduous and contingent process, which for the West had happened to work out well, at least in some crucial time periods.

In the case of Huntington, the hierarchical difference is drawn along more innate, cultural lines, where people of different civilizations are ascribed culturally deterministic identities. As briefly outlined above, the critical discourse that emerged in response to Huntington's theory of closed and mutually exclusive civilizations has in powerful ways – both empirical and theoretical – refuted many of his assumptions, or at least shed critical light on it. It is for this reason that this rich discourse, joined not only by postcolonial voices but also realist IR scholars, poses a serious challenge to Huntington's theory. Yet despite the persuasiveness of this academic discourse, one cannot fail to notice how Huntington's specter of a clash of civilizations and its implicit threat to "Western" values has successfully penetrated political and public discourse, and is only gaining widespread popularity amongst the far-right movement in Europe, as well as in the United States.

In Germany, for example, the right-wing party *Alternative für Deutschland* [trans. 'Alternative for Germany'], also known as "AFD", was extremely successful at the state and federal electoral level since its inception in 2013, running a campaign that was chiefly focused on eurosceptism, anti-immigrant sentiment, the fear of Islamic religious values clashing with 'European' ones, as well as anxieties about the demographic demise of White Germans[38]. It is important to note that the AFD was founded by a highly educated elite from within German academia and cultural-political circles, and that its most prominent founder, Bernd Lucke[39], is a professor of macroeconomics at the University of Hamburg and former World Bank advisor. Rallying around the 'battle of civilizations' has since become popular amongst AFD supporters, both for activists on the streets and journalists and intellectuals who publish on this topic[40]. The main thrust there, similar to Huntington, is the antagonistic and hierarchical pitting of homogenous cultural groups against each other, and the sense that Western civilization is absolutely incommensurable with any other value system – which is why, in their opinion, Muslims and other minority groups who live in the West have to leave.

The phenomenon of right-wing parties' success and their attempts of running their platforms on the vision of civilizationary clash is no way restricted to Germany alone, nor

[38] The full AFD party manifesto can be found here: https://www.alternativefuer.de/wpcontent/uploads/sites/7/2016/05/2016-06-27_afd-grundsatzprogramm_web-version.pdf. For the most comprehensive political analysis of the rise and success of the AFD see Alexander Häusler, ed. (2016) *Die Alternative für Deutschland: Programmatik, Entwicklung und Politische Verortung*. Wiesbaden: Springer Verlag.

[39] Barbara Supp, "A portrait of Bernd Lucke and the new German Right", *Der Spiegel Online*, 22.05.2014. At: http://www.spiegel.de/international/germany/a-portrait-of-bernd-lucke-and-the-new-german-right-a-969589.html

[40] Heinz Theisen, "Im Kampf um die Zivilisation" [The battle for civilization], Der Tagesspiegel, 10.05.2015. At: http://www.tagesspiegel.de/politik/islam-und-islamismus-im-kampf-um-die zivilisation/11755982.html#commentInput. This article was recommended by the AFD party on their website: http://www.afd-ofl.de/leseempfehlung-im-kampf-um-die-zivilisation/.

is it an anomaly in the electoral trends of post-Cold War Europe: in recent elections, the staggering success of right-wing parties in France, Austria, Sweden, Denmark, the Netherlands, Italy, Greece and the UK, have baffled political analysts and mainstream-party politicians alike (Akkerman, De Lange, & Rooduijn, 2016; Hartleb, 2011). Likewise, in the US, as can be seen in the 2016 presidential election[41], the fear of a beleaguered Western civilization under threat is potent and alive. The roots of this might go back to President George W. Bush (2001), when he described the new enemy four days after the September 11th attacks as "a group of barbarians [who] have declared war on the American people" – the first time he used the 'barbarian' term.

The aim of this chapter is not to question the clashing civilizations theory using historical or theoretical approaches that focus on refuting the intellectual and moral *content* of the theory. Rather, this chapter will examine how the cognitive *mechanisms* involved in the in-group/out-group positioning and the hierarchical ordering of different human groups within the clashing civilizations theory are problematic for cooperative politics. The choice of cognitive mechanisms over ontology is a deliberative one within this dissertation, and runs as a thread through all the chapters.

Within this chapter, I use the PBP to criticize the clashing civilizations theory based on its rigid and nonnegotiable demarcation between different civilizational groups, as well as the inherent sense of the superiority of Western civilization in Huntington's position, arguing that both aspects lead to a potentially dehumanized perception towards out-groups at the brain level. This, as shown in earlier chapters, is associated with de-

[41] Ishaan Tharoor, "Donald Trump's real foreign policy: a clash of civilizations", *The Washington Post*, 24.04.2016. At: https://www.washingtonpost.com/news/worldviews/wp/2016/04/28/donald-trumps-real-foreign-policy-a-clash-of-civilizations/

mentalizing, disgust and questionable moral judgments towards out-groups. I will proceed to mainly advancing a theoretical argument, but also discuss recent studies coming from Emile Bruneau's (2015; 2016) lab at MIT, in which he discovered troubling links between rating different groups along a civilizationary scale, dehumanization rhetoric and inter-group violence. I will build on the theoretical claims laid down in earlier chapters and focus in this chapter on applying my claims at the International level, particularly to the case of the Bosnian genocide, and within discourses in the IR field.

The main point I wish to bring across is that in order to effectively criticize the historical, political and moral pitfalls inherent in the civilizationary clash theory, the PBP on dehumanized perception is crucial in highlighting the cognitive dangers in this kind of thinking about one's own in-group and global politics at large. Unlike other more conventional avenues of criticism, the PBP allows us to consider what effect civilizationary clash thinking might have on our mentalizing and humanizing capacity of other out-groups, and take into account attitudinal and behavioral outcomes such as aggression, violence or neglect.

Although postcolonial and historical criticisms of the civilizationary clash theory can prove powerful in intellectually rejecting Huntington's worldview, they cannot explain *why his theory is cognitively problematic* and thus fail to understand why it could be so compelling for the brains of the people who operate on this kind of worldview. Furthermore, they fail to explain why up until today, his theory remains highly compelling for elites and ordinary citizens alike, such as for example the elite founders and cross-class supporters of the German AFD party. This is what this chapter sets out to do.

5.2. Ambiguous Categories of 'Civilized' and 'Barbarians' in History

I argue that the 'clashing civilizations' theory is powerful and enduring not because of its intellectual sophistication or historical accuracy, but because *it exploits in a fundamental way how our brains operate in the social world*, i.e. through the construction of in-group identities vs. out-group ones (Tajfel and Turner 1986; Brewer and Miller, 1996; Dovidio, Gaertner, Kawakami, and Hodson, 2002; Hewstone, Rubin, and Willis, 2002). Indeed, Huntington's own definition of civilization goes straight to the heart of this: "Civilizations", he says revealingly, "are the biggest "we" within which we feel culturally at home as distinguished from all the other "thems" out there" (ibid., 1996: 43). Just as with any in-group bias, the civilizationary in-group belonging depicted by Huntington serves "not just [as] a matter of rational self-interests but may also include more symbolic and emotional benefits to the group" (Spears, 2007: 484). However, as inevitable and even natural in-group identification might be, I argue that the civilizationary pitting of "us" against "others" is one of the most exclusionary and problematic kind of in-group belongings – because it involves a hierarchical and antagonistic positioning of the *superior civilized* in-group vs. the *inferior barbaric* out-group, potentially facilitating dehumanization at the cognitive level, and as a result, neglect and aggression at the behavioral one.

But what about past empires and their pitting of the civilized insider against the barbaric outsider, one could ask? Was this kind of divisive distinction not commonplace in various parts of the world, throughout history? The Ancient Greeks, who coined the term first with Homer and his description of the Carians as '*barbarophonoi*', viewed

anyone who could not speak the Greek language as barbaric – i.e. membership of Greek in-group identity was chiefly constructed along linguistic and political participatory lines (Bacon, 1961; Vlassopoulos, 2013). It is important to note here that the Ancient Greeks were themselves confused, at least until the time of Alexander the Great, who actually spoke Greek, since many local and regional dialects were mutually comprehensible and cross-pollinated linguistically (Vlassopoulos, 2013).

In the case of the Roman Empire, barbarians were even more loosely defined – the word served as an umbrella term for all foreigners and those people who encroached the imperial borders. Historians believe that Romans "conceptualized groups of barbarians not for their specific traits but for their collective appearances with other groups of barbarians" (Mathisen, 2016: 17), suggesting that the term was used more as a descriptive rather than trait-specific or derogatory term that targeted a particular out-group.

Further to this, the demarcation between civilized Romans and foreign barbarians was permeable in the sense that foreigners could change allegiances and transform themselves to the point of being absorbed by Roman identity (Ladner, 1976; Goffart, 2006). Indeed, the latest scholarship on the relationship between Romans and barbarians suggests that it was marked by interconnectedness, dynamic exchange and fluctuation, going so far as to state that "Romans of Late Antiquity did not have "barbarians on their mind" and were not obsessed with "barbarophobia" to nearly the extent that modern commentators seem to think" (Shanzer and Mathisen, 2016: 2). Even though it is important not to forget that Romans of Late Antiquity, including Christians such as Augustine, harbored "stereotype[s] of the bestial barbarian who must be tamed as wild

animals are tamed: overpowered then calmed, for they cannot be persuaded by reason" (Clark, 2016: 35), the enacted relationship between 'civilized' Romans and 'barbarian' foreigners was undeniably a reciprocal one, where cultural interaction and economic exchange, rather than hostility, dominated everyday life in the empire.

Similarly, in the context of the Chinese Empire, where modern onlookers commonly hold the notion that the Northern frontiers of the empire were marked by "a set of dual oppositions – between pastoral and settled people, between nomadic tribes and Chinese states, between an urban civilization and a warlike uncivilized society" (Di Cosmo, 2002: 2), whereas in fact, "a single term analogous to the European barbarian did not exist in Ancient China" (ibid.: 7)[42] and that in fact, just as in the case of the Roman Empire, insiders and outsiders of the empire engaged in a variety of exchanges, interactions and alliances (So and Bunker, 1995). Revisionist Chinese historians who refuse to rely solely on texts and rhetorical analysis to understand the relationship Chinese had with outsiders, but consider instead the context of political relations within the empire itself, point out that the stark distinction drawn by rulers between the cultural unity of the Chinese "Hua-Hsia" and external barbarism was in fact a strategy to unite factional states within China itself (Di Cosmo, 2002: 7). In other words, instead of being driven purely by "barbarophobia", rhetoric about the civilized vs. the barbaric world in Ancient China often had expedient causes originating from domestic political instabilities.

[42] Sima Qian, China's first major historiographer, who composed his masterpiece *Records of the Grand Historian* (1971) around 100 B.C., first introduced the idea of the 'barbaric' other in the form of the steppe nomadic people of the North, whom he called "Hsiung Nu" (匈奴). In Ancient China, various names existed for the Northern and Western frontier peoples, such as "Dong Yi" (東夷) and "Xi Rong" (西戎), amongst others.

A completely different perspective comes from the beginnings of the Ottoman Empire. As much as the siege and sack of Constantinople by the Turks in 1453 was a gruesome affair, historians such as Margaret Meserve question whether the eyewitness reports and subsequent recollections of the Turks' brutish and quasi-animalistic rampage during the conquest were not somewhat exaggerated, building in fact on "a long tradition of Christian rhetoric, dating back at least to the First Crusade, had aimed at *dehumanizing* the Muslim foe [my emphasis]" (Meserve 2009: 66; Tolan, 2002). Renaissance humanists, in fact, hoped to use the rhetoric of the brutal Turkish 'barbarian' in order to persuade their compatriots of the necessity of a new crusade, leading Meserve to conclude that "the motivation for attacking the Turks on so many fronts was rhetorical" (ibid.: 67).

However, the impression that the Turks gave of themselves and their culture was anything but 'barbaric': their military and diplomatic skills, the willing obedience the sultan managed to command from his subjects, as well as the intellectual sophistication of Islamic religion and its related artistic and cultural outpourings were awe-inspiring to many outside contemporaries and observers (Tolan, Laurens, & Weinstein, 2012). In other words, Renaissance humanists managed to conveniently and successfully dehumanize Turks as animal-like barbarians for their own expedient political purposes, but at the same time, the overall response to the Turks political and cultural achievements was much more ambiguous and conflicted, if not marked by a certain level of respect and admiration amongst Christians.

The upshot of these historical examples is that dividing the world into civilized in-groups and barbarian out-groups is widespread in world history. One could argue that

one common reason to rally around a civilizationary in-group identity in the above cases is a *sense of threat*: "although a sense of "belonging" to the community might exist prior to an external challenge, the fact of being challenged makes its members acutely aware of their common boundaries" (Di Cosmo, 2002: 2). I will return to the connection between in-group threat and dehumanization of out-groups when I discuss the Bosnian genocide.

Boundaries between civilized "us" versus barbarian "thems", as Huntington himself puts it, are by far less bounded, rigid and incommensurable as in his clash of civilizations theory. Throughout the history of different empires, the civilized in-group and the barbarian outsiders communicated, influenced and fundamentally changed each other's identities. The reason for erecting a division between "us" and "them" were less due to inherent, essence-like differences, but often served politically expedient reasons such as domestic instability and a perceived sense of threat from the outside. By stressing this, I side with postcolonial and historical critics of Huntington's theory who believe that his theory is misguided on various historical, empirical and theoretical grounds. Martin Hall and Patrick Thaddeus Jackson (2007) sum it up best when they, in opposition to Huntington, postulate that:

1. Civilizations are weak, not bounded.
2. They are loosely integrated.
3. They are heterarchical, not centralized.
4. They are contested, not consensual, in that power struggles over material and symbolic resources, and disputes over meaning and purpose abound, and finally
5. Civilizations are in a state of flux; they are processes and relations (p. 7-10).

Yet as powerful as Huntington's critics are in deconstructing some of his misguided ideas about how civilizations function and why antagonistic relationships with outsiders might arise, I believe that these critics have failed to respond to an important

challenge and uncomfortable truth that Huntington's theory reveals, namely, what if the *differentiating between a superior "us" versus an inferior "them" is actually how our brains function* in social settings? By looking at the historical evidence alone, it is difficult to deny that this tendency exists across cultures, geographies and time periods.

The problem with Huntington however, I argue, is that he is unaware of the cognitive mechanisms at play in extreme in-group/out-group perception and the extent to which a 'barbaric' view of an out-group can lead to dehumanizing and de-mentalizing them. Huntington and those who borrow his civilizationary clash theory do not realize that by exploring our inherent tendency to divide our social world into "us" and "them" to the very extreme – i.e. amplifying our perception of in-group belonging to the point that any relationship with an out-group is anticipated as a battle or clash – they invite our brains to exclude and de-mentalize other out-groups to an extent where intractable and violent conflicts become more likely.

As the above historical examples have shown, even though our social brains might function in terms of in-group/out-group distinction, they are also equipped to overcome some of these biases and divisions for the sake of inter-group communication, exchange and cooperation. Even though the Ancient Greek, Roman, Chinese and Ottoman Empires were all defined by distinguishing themselves from foreigners and barbarians, the relationship with barbarians was not exclusively marked by isolation and antagonism, but instead, the civilized and the so-called barbarians often engaged in reciprocal economic, cultural and political relationships, resulting in a deep interconnectedness, as well as cross-pollinating influences on each other's own identity and self-conception. History, therefore, unlike the civilizationary clash theory, paints a

much more ambiguous, nuanced and creative picture in respect to how we can challenge and transform our innate cognitive tendencies. In order to do so effectively, however, we need to understand which exact cognitive mechanisms are at stake.

The point of this chapter is to show that no matter how forecful a historical argument might be in refuting the civilizationary clash theory, we still need the PBP in order to access the cognitive mechanisms that Huntington and his followers tap into and amplify. Instead of rejecting the civilizationary clash theory as outright wrong, I wish to address the fact that in-group/out-group divisions and biases do (and will continue to) exist and are fundamentally part of how our social brains function. However, unlike Huntington, this dissertation contends that the current research evidence in prejudice and in-group bias reduction coming out of the social psychology and neuroscience fields show that we have reasons to be hopeful that these divisions can also to an extent be overcome.

5.3. Why the PBP on Civilization Theories Matters

5.3.1. Why the PBP on Civilization Theories Matters I: The Extreme Effects of Subtle Dehumanization

It is undeniable that to most of us, dividing one's world into groups that one feels one belongs to and other groups that one views as outside or foreign to one's sense of identity, seems completely natural. Feeling part of a cultural, gender, occupational or religious in-group, but also more specialized in-groups such as one's local bird watching or dancing club, or a charity or civic activism society, can feel rewarding and empowering. Indeed, Robert D. Putnam (1993) famously argued that local civic group

memberships in the Italian North versus a lack thereof in the country's South led to more favorable conditions for democratic institutions in the North[43], effectively designating a highly positive role for civic in-group memberships in the development of democratic accountability and sustainability. However, when we think of in-group identification, out-group exclusion in the form of prejudice, discrimination and racism also come to mind, suggesting a darker side to the naturalness and inevitability of in-group belonging. It seems that as much as in-group identification can contribute to our nobler civic aspirations, it can also become a breeding ground for the most disruptive chauvinistic and tribal tendencies in society. But is the need to identify with an in-group really inevitable? And does in-group preference automatically entail out-group prejudice?

If we look at the psychological (and increasingly, neurocognitive) research into in-group belonging, a more complicated picture emerges. Understanding in-group favoritism and how it might lead to interpersonal conflict and discrimination has been one of the major research questions for social psychologists in the 20th century, especially after extreme manifestations of nationalism and racial superiority in the two World Wars. One group of psychologists, led by Muzafer Sherif, focused on how competition over scarce resources and conflict over materialistic goods determined in-group bias (Sherif, Harvey, White, Hood, and Sherif, 1961). As a reaction against what Sherif termed his "realistic conflict theory", Henri Tajfel and his colleagues created the "minimal group paradigm"[44], in which they postulated that not all inter-group discrimination is due to competition over resources (Tajfel, 1970). Instead, Tajfel discovered that minimal

[43] For critical reviews of Putnam's controversial correlation between civic community memberships and the efficiency of democratic institutions see Tarrow (1996) and Barceló (2014).
[44] See Tajfel (1982) for his own critical review of the theory.

conditions, such as being randomly assigned to a group that favors one painter over another, were sufficient for strong in-group identification and the wish to allocate positive rewards to one's in-group – though it has to be stressed here that this in-group bias did not necessarily entail out-group discrimination or the wish to do them harm (Mummendey & Otten, 1998).

What the minimal group paradigm shows is that neither materialistic rewards nor impending conflict as such were necessary for in-group alliances and bias to occur – eventually, this would lead to the development of "social identity theory" (Tajfel and Turner, 1986; Hogg and Abrams, 1988), which explored why members of a group discriminate in favor of their members because of symbolic and psychological, rather than materialistic rewards. The upshot of this is that in-group bias does not always have to be motivated by blatant self-interest, such as the securing of resources, but might occur for non-materialist, more subtle reasons as well.

Another level of complexity that we need to consider is the automatic and unconscious nature of some in-group biases. The "Implicit Association Test" (IAT), developed by Greenwald, Banaji, McGhee and Schwartz (1998) takes into account that much of the out-group discrimination that might result from in-group bias might in fact not be accessible to us in explicit self-reporting. In other words, much inter-group discrimination might often not be detectable if we simply asked people how they felt about certain out-groups.

For example, in the case of race-based discrimination, the phenomenon of 'aversive racism' reflects the subconscious nature of our out-group prejudices (Dovidio & Gaertner, 2005). Unlike "old-fashioned racism", aversive racism "represents a subtle,

often unintentional form of bias that characterizes many White Americans who possess strong egalitarian values and believe that they are nonprejudiced" (Dovidio, Gaertner, Kawakami, and Hodson, 2002, p.90). People who engage in aversive racism face a genuine conflict between denying subjective prejudice and yet having underlying negative feelings and beliefs towards another racial out-group; therefore the manifestation is implicit, not explicit racial prejudice.

For example, aversive racism in an everyday situation such as a hiring process might manifest in the following way: when White candidates were asked to evaluate the applications of White and Black applicants who were equally qualified for the job, there was no discrimination against the Black applicant. However, when the candidates' qualifications were more ambiguous or problematic, White candidates chose the Black applicant significantly less often than a White candidate with the same credentials (Dovidio & Gaertner, 2000). This study suggests that aversive White racists are more willing to give White applicants the "benefit of the doubt", but which they are not willing to extend to Blacks. Alternatively, an aversive racist's feelings can be described as "more diffuse, such as feelings of anxiety and uneasiness" (Dovidio et al. 2002: 90). It is important to note here though that aversive racism is *not* viewed as a kind of psychopatholoy, but instead as rooted in ordinary and also adaptive processes (Dovidio and Gaertner, 1998). Similar to subtle dehumanization and infra-humanization, which I have discussed in earlier chapters, implicit out-group prejudice is something that we all experience, including minorities themselves (Sidanius and Pratto, 1999). Subtle out-group discrimination should therefore not be vilified or pushed to the margins of the

political debate, but in order to tackle it, needs to be addressed and acknowledged as a widespread phenomenon and *natural brain capability* in the first place.

I argue further that another kind of closely related out-group discrimination, infra-humanization, can become particularly problematic in the context of the civilizationary binary of "us" between "them". As a reminder, infra-humanization is the process in which we only ascribe 'secondary', i.e. distinctly human, emotions to members of our in-group, whereas we ascribe solely "primary emotions" to out-groups and animals (Leyens et al., 2001). Put differently, we attribute an exclusive *psychological essence* to members of our in-group, in so far as we believe that only in-group members have emotions that are uniquely associated with humans, such as complex feelings of pride, shame, hope, hate and despair, for example (Leyens et al., 2002). Out-groups are denied that human uniqueness at the emotional level and are therefore considered less uniquely human than "us". Even when secondary, uniquely human emotions are invoked for an out-group, we will only activate a "humanity concept", i.e. the idea that someone belongs to humanity, for in-group members (Vaes, Paladino, and Leyens, 2006). It is important to remember here that infra-humanization, just like aversive racism or implicit bias, is a subtle process that we might not be fully aware of (or admit to) when interacting with our own in-group and other out-groups.

However, just because infra-humanization of one's in-group happens implicitly does not mean that there are no considerable consequences at the behavioral level. For example, in response to secondary emotions by a fellow in-group member, we are more willing to help them, show increased perspective taking and imitation than when these same secondary emotions are displayed by an out-group individual (Vaes, Paladino,

Giovanazzi, Castelli, & Leyens, 2003). In addition, other attitudinal consequences of infra-humanization are lack of forgiveness for the out-group and justification (rather than guilt) for past misdeeds committed by the in-group against the out-group (Castano, Giner-Sorolla, & Leyens, 2006; Demoulin et al., 2007).

If applied to the context of political campaigning in Italy, political slogans that employ secondary emotions only seemed to be effective in commanding conformity if viewers identified the campaigning politician as part of their in-group (Vaes, Paladino, & Magagnotti, 2011). If participants in this study were presented with a political slogan employing primary emotions (e.g. fear or anger as in "*With this government, the future makes us afraid*"), no bias in preference for an in-group or out-group political candidate was found, whereas if slogans used secondary emotions (e.g. shame or pride as in "*Shame to this government: salaries are in lire* [Italy's previous national currency], *prices are in Euro* [Italy's current currency as a member of the European Union]"), participants were much more inclined to stand behind the political opinion of their own in-group candidate rather than an out-group one.

This study therefore shows that infra-humanization, triggered by the usage of uniquely human emotions for one's in-group, can be a powerful and decisive factor in people's support of political opinions, even if the psychological processes at play might be manifested in subtle and implicit ways. Infra-humanization also plays a role in perceiving someone as a symbolic threat to one's in-group's welfare: a study on Portuguese infra-humanization of Turkish people showed that it led to the perception of them as a symbolic threat and thus predicted opposition to include Turkey in the European Union as a new member (Pereira, Vala, & Leyens, 2009). At its core, the idea

of infra-humanization is based on the primacy of the in-group and might be "the only way to distinguish between groups" (Leyens et al., 2003, p.710) when the existence of political taboos and standards of political correctness *forbid the explicit expression of nationalist or racial prejudice*. Most crucially, infra-humanization of out-groups withholds a human essence from them, by denying them emotions that would normally distinguish humans from animals.

In earlier chapters, I discussed another manifestation of subtle dehumanization, namely the dual model of dehumanization, where groups are dehumanized along *animalistic* and *mechanistic* lines (Haslam, 2006; Haslam & Loughnan, 2014). In the case of animalistic dehumanization, out-groups are denied *uniquely human* (UH) traits that distinguish humans from animals (e.g. cognitive aptitude, refinement, and civility), whereas for mechanistic dehumanization, out-groups are denied *human nature* (HN) traits (e.g. warmth, individualism, creativity, and emotionality). Studies on the real-life political behavior and attitudes showed significant effects of both kinds of subtle dehumanization.

For example, when testing Italians' willingness to help Haitian and Japanese earthquake victims, a study found that Italians animalistically (but not mechanistically) dehumanized Haitians and that they mechanistically (but not animalistically) dehumanized Japanese – both of which led to decreased willingness to help either group of earthquake victims (Andrighetto, Baldissari, Lattanzio, Loughnan, & Volpato, 2014). In another study, mutual mechanistic dehumanization between Palestinian and Jewish Israeli people predicted preference for punitive forms of justice over restorative forms of justice (Leidner, Castano, & Ginges, 2013).

In an extremely troubling study, Viki and colleagues (2013) found that Christians associated more animal-related words than human-related words with Muslims, which in turn predicted their support for torture of Muslim prisoners. In short, both infra-humanization and the Haslam's dual model of humanization, though highlighting rather subtle and oftentimes seemingly innocuous or benign ways in which we deny other out-groups their humanness, can nonetheless have destructive political effects in a variety of interpersonal conflict, punitive, and humanitarian aid situations.

If we apply these psychological insights to the political debate about the 'civilizationary clash' theory, the importance of paying attention to subtle dehumanization of out-groups is undeniable. Edward Said (2001), in his article "The Clash of Ignorance", addressed Huntington's prediction that the two civilizations that were most likely to come into conflict were Islam and Western Judeo-Christian culture. Said is appalled at the historical ignorance and analytical over-generalization in Huntington's theory, and points instead to the interconnectedness between the West and Islam, and the plural voices and diverse internal developments within Islam itself.

Writing in the immediate aftermath of the September 11th attacks, where George W. Bush (2001) described the attackers as "a group of barbarians [who] have declared war on the American people", and an overwhelming majority of public intellectuals and journalists jumped eagerly on Huntington's 'civilizationary clash' bandwagon to explain the significance and causes behind the attacks, Said views the 'civilizationary clash' theory as a "gimmick", which in troubling ways reinforces a sense of "defensive self-pride" and "gigantism and apocalypse" amongst Western nations (ibid., p.12). In a subsequent publication, Said is even more blunt in his criticism of Huntington, calling the

clash of civilizations thesis the "purest invidious racism, a sort of parody of Hitlerian science directed today against Arabs and Muslims" (Said, 2004, p.293).

One could construe Said's analysis and subsequent comments as exaggerated and paranoid – what, after all, is so problematic about Huntington's pitting of the Western "us" against the non-Western "thems"? How can this possibly amount to racism? Based on the above evidence, we can counter to this that as aversive racism and infra-humanization show, there exist powerful indirect and implicit ways to express out-group derogation and in-group superiority, without having to make explicit racist statements as such.

Said's critique therefore could be strengthened by the PBP and psychological evidence on implicit in-group favoritism, since his historical line of argument is unable to specify and address the exact mechanism of subtle in-group humanization inherent in Huntington's thesis. That being said, this chapter disagrees with Said and other critics that Huntington is committing an empirical or theoretical offense by dividing our political world into cultural in-groups and out-groups, since his assumptions about the naturalness of in-group identification are not completely implausible from a psychological viewpoint. Rather, where Huntington goes wrong is in the dangerous amplification and radicalization of a 'Western' in-group identity, to the point of presenting Western civilizationary identity as so completely unique and superior to backward Islamic culture that the *ascription of equal humanness in regard to other non-Western identity groups* becomes impossible.

I argue that the civilizationary clash theory is dangerous for international peace not only because it misrepresents the interconnected relationship between different

cultural and religious groups in history, but because it engages in subtle in-group humanization rhetoric such as infra-humanization, which nevertheless can result in negative behavioral and attitudinal consequences, such as lack of concern and forgiveness for the out-group, unwillingness to accept responsibility for past misdeeds against the out-group, and a general inability of perspective taking on behalf of the out-group. On the latter point in particular I have written in chapter 3 in regard to *mentalizing*, which is a fundamental kind of perspective taking that we develop cross-culturally in infancy. In other words, one could argue that the 'civilizationary clash' theory makes it harder to mentalize other cultural groups, especially in Huntington's version, where different groups are treated as isolated and bounded units that are divided by essence-like and incommensurable differences.

Isaiah Berlin once made a helpful distinction in regard to the conception of civilization, outlining two approaches: relativist and pluralist. He described the relativist approach as follows:

> "the most extreme versions of cultural relativism, which stress the vast differences of cultures, hold that one culture can scarcely begin to understand what other civilisations lived by – **can only describe their behavior but not its purpose or meaning**, as some early anthropologists described the behavior of savage societies [emphasis added]. If this were true (as, for example, Spengler, and at some moments even Dilthey, seemed to say) the very idea of the history of civilisations becomes an insoluble puzzle"
>
> (Berlin, 2013, p.84f.)

Based on Berlin's assessment, Huntington is a civilizationary relativist as well, not only because of his indebtedness to the work of Oswald Spengler, but most important, because of his belief that the West will never be able to fathom the 'purpose and meaning' of other non-Western civilizations (and vice versa), and therefore should refrain from engaging with them. Although we could potentially interpret this kind of relativism as a

respectful acknowledgment of irreconcilable differences, Berlin suggests that there is a darker aspect to this: other civilizations are viewed as so completely different and incomprehensible to us that they are reduced to the category of 'savages' (or alternatively 'barbarians') – animal-like creatures with primary but no secondary, uniquely human emotions. Berlin then presents and advocates for another kind of approach, the pluralist one:

> "(…) the values of these remote peoples are such as **human beings like ourselves** – creatures capable of conscious intellectual and moral discrimination – could live by. These values may attract or repel us: but to understand a past culture is to understand **how men like ourselves**, in a particular natural or man-made environment, could embody them in their activities, and why; by dint of enough historical investigation and **imaginative sympathy**, to see how human (that is, intelligible) lives could be lived by pursuing them [emphasis added]."
>
> (ibid., p.86)

It is remarkable how Berlin stresses and details the ability to view other cultural groups through the lens of our shared humanity, as well as mentioning the need for imaginative sympathy – i.e. sympathy being the highest form of empathy, which builds directly on our universal mentalizing abilities (McCall and Singer, 2013). It is worth noting here that Berlin's remarks could be interpreted as well-meaning but naïve, in that his demand for a shared sense of humanity when imagining other cultural out-groups would be nice if met, but really is not central to solving real-world clashes between 'civilizations'. Here again the PBP and experimental evidence on the infra-humanization of out-groups (as well as how it affects actual decisions and behaviors) become crucial for strengthening Berlin's theoretical argument.

Another point that Berlin is making in this quote in favor of the "pluralist" civilizationary outlook is more subtle: by asking us to humanize cultural out-groups, he is also asking us to imagine them as fully human individuals, whose life choices might not

align with our own, but which were nonetheless taken by *someone* and not some homogenous, unintelligible group. I mentioned in previous chapters the effective role of individualizing and individuating people who belong to an out-group, which can help in overcoming and preventing dehumanized perception at the cognitive and brain level (Swencionis & Fiske, 2014).

Echoing Berlin, the psychologists studying infra-humanization and subtle dehumanization state in their policy recommendations that "to combat infra-humanization, rather than emphasizing differences and similarities between groups, politicians, media, and educators should insist upon complementarities and universalism" (Leyens, Vaes, Gaunt, & Paladino, 2007: 160) as well as "emphasizing understanding, accepting, and showing concern for the welfare of all human beings, even those whose life differs from one's own" (Roccas, Klar, and Liviatan, 2006, p.137).

In conclusion, it might be worth reflecting on another historical example here. In trying to explain the decisive reason that brought about the formation of NATO, Patrick T. Jackson (2003) suggests that rhetorical appeals to a shared 'Western civilization' between the U.S. and Europe – in contrast to the Soviet Union – played a decisive role in NATO's eventual formation. For Jackson, the U.S. needed to be convinced to join a British-led NATO initiative, which the latter achieved by employing "occidentalist" language of Western exceptionalism and the rhetorical "nesting" of NATO states within an imagined community of shared Western civilization (ibid.: 245).

Instead of proposing realist, liberal or constructivist explanations of this crucial moment in NATO's history, Jackson offers what he calls a "relationist" approach, which neither locates causal mechanisms exclusively at the individual level (i.e. realism and

liberalism) or at the level of social totalities and systems (i.e. constructivism) but at the place where "patterns of social practice" determine political outcomes. In other words, contra explanations of NATO formation that assume that NATO came into existence predominantly because it was needed as a defensive alliance between individual states, Jackson argues that the civilizationary arguments and pitting of a Western "us" against a hostile and foreign Soviet "them" created a sense of common in-group identity, which in turn became such a powerful narrative that it created sufficient consensus on both sides of the Atlantic.

Given our insights into the deep-seated psychological and cognitive need for in-group belonging, and the way how the civilizationary argument feeds and amplifies this need, Jackson's alternative explanation of the NATO formation process and his highlighting of the significance of civilizationary rhetoric in a climate of geopolitical uncertainty is worth considering.

5.3.2. Why the PBP on Civilization Theories Matters II: Blatant Dehumanization and Intergroup Violence

To recap the argument, the subtle forms of dehumanization discussed above have one thing in common – they are not just an expression of dislike for an out-group (as for example prejudice is), but they are an indirect and implicit expression of our sense that someone else does not count as a full human being to us. This sense is distinct from disliking someone in that the exclusion inherent in subtle dehumanized perception is both more profound and yet less straightforward than mere dislike: it is more profound because being unable to include someone else in my sense of humanity carries more

philosophical and moral weight than the feeling of 'I don't like you for X reason'; it is less straightforward because subtle dehumanization is measured along trait attributions (animalistic, mechanistic), the denial of secondary emotions and disgust, instead of explicit dislike reasons.

The field of subtle dehumanization study emerged over a decade ago in response to the existing field of explicit dehumanization (Kelman, 1973; Bandura et al. 1975; Opotow, 1990). The field arose in the aftermath of grappling with WWII's atrocities and genocide, thus locating the occurrence of dehumanization in situations of extreme hostility and violence. Subtle dehumanization, on the other hand, is thought to occur in everyday situations and to be experienced by the dehumanizer in often unconscious and automatic ways. I argued that despite its subtle manifestations, this kind of dehumanization can be politically problematic when employed in the context of the civilizationary clash theories, where inferior civilizations are often described in brutish, savage and barbarian and denied uniquely human emotions. It is possible that the analytical tool of subtle dehumanization is convenient in the academic context, where there exist standards of political correctness and linguistic propriety that forbid an explicit expression of disgust or contempt for an out-group, but are instead expressed implicitly and even unconsciously. This being said, however, racism, sexism or tribalism are nonetheless still known to exist and enacted in seminar rooms, lecture halls and between the lines of academic research books and articles (Berg, 2009; Coleman, 2005; Patton, 2004).

What happens however when in the real political world, where viciously fought intergroup conflicts and violent clashes do not adhere to any such standards of political

correctness, and as an effect, political actors use much more blunt and direct language to express their disdain for another group? One only has to think of how Jews were portrayed in Nazi propaganda as rats, African-Americans as apes during slavery, Tutsis as cockroaches, and Romani people as vermin by Europeans (Goff, Eberhardt, Williams, and Jackson, 2008). More recently and ongoing, in the Israeli-Palestine conflict for example, Palestinians are described as "wild beasts"[45] and Israelis as "killing machines"[46]; whereas in the Mediterranean migrant crisis in the migrants have been called "cockroaches, "swarms", "brutes" and "scum", by both European right-wing and mainstream conservative politicians alike[47]. Other examples are soccer fans who throw bananas at black football players in Europe, or President Obama being depicted as an ape in political cartoons[48]. It is for this reason that a new line of research has turned its attention to so-called *blatant dehumanization*, which involves explicit beliefs about the biological and human inferiority of certain out-groups (Kteily, Bruneau, Waytz, & Cotterill, 2015). With blatant dehumanization, people do not so much indirectly associate certain groups with non-human traits, or deny them uniquely human traits, but instead differentiate very openly between their in-group and others – in a hierarchical fashion where other out-groups are inherently inferior to one's own in-group.

In order to measure blatant dehumanization, Emile Bruneau and his colleagues devised an "Ascent measure of blatant dehumanization", which consisted in an "Ascent

---

[45] http://newobserveronline.com/israel-builds-fence-against-wild-beasts/

[46]http://www.israeltoday.co.il/NewsItem/tabid/178/nid/24151/Default.aspx?topic=article_title

[47] https://www.theguardian.com/uk-news/2015/jul/30/david-cameron-migrant-swarm-language-condemned; http://www.zeit.de/gesellschaft/zeitgeschehen/2016-04/lutz-bachmann-pegida-gruender-volksverhetzung-gericht

[48] Huffington Post. (2014, March 24). *Belgian newspaper accused of racism for picture of Obama and Michelle as apes*. Retrieved from http://www .huffingtonpost.com/2014/03/24/newspaper-obama-ape-belgian-satire- putin-barack-president-racism-racist_n_5020987.html/

of Man" diagram that depicted five different figures, from an ape-like animal to an up-right standing, human-looking person. Underneath the diagram is a list of different national and religious groups (American, Canadian, European, Chinese, South Korean, Muslims, etc.) with sliders. The instructions accompanying the diagram read as follows: "People can vary in how human-like they seem. Some people seem highly evolved whereas others seem no different than lower animals. Using the image below, indicate using the sliders how evolved you consider the average member of each group to be" (ibid.: 904). Since Bruneau's study wanted to test the potential blatant dehumanization of Arabs and Muslims by Americans in particular, they recruited American participants for this study and excluded minority groups that were listed next to the sliders (such as Asians, Latinos/Hispanics and Middle Easterners/Arabs). The results showed that Europeans, Canadians and Japanese were rated as similarly evolved as Americans, whereas South Koreans, Chinese and Mexicans were deemed as significantly less evolved. The lowest on the scale, however, were Arabs and Muslims (i.e. rated as resembling the ape-like figure at the lowest end of the "Ascent of Man" scale). In addition, they found that a measure called "social dominance orientation", which reflects an active orientation towards enforcing hierarchy between groups, was strongly associated with blatant dehumanization.

In the context of this same study, Bruneau and colleagues the further tested how blatant dehumanization would predict support for various policies and decisions towards out-groups. They found that blatant dehumanization "predicted support for minimizing Arab immigration, less compassionate responses to injustice experienced by an Arab target, and less money donated to an Arab versus American cause" (ibid.: 910). In

addition, subtle dehumanization was *not* a significant predictor for various important political decisions, which highlights the need to acknowledge blatant dehumanization as a unique way of excluding an out-group from one's sense of humanity in its own right. Similarly, prejudice defined as disliking another group did *not* determine the outcome of immigration support for or objection to injustice done to Arabs. Thus the common wisdom that disliking someone must surely be a decisive indicator for intergroup conflict might not necessarily be true – politically we therefore need not only to pay attention to how certain groups are disliked, but also, how they are viewed and portrayed in a hierarchically dehumanizing and animalistic way.

So far, we have gained an understanding of the perspective of the dehumanizer and how both subtle and blatant dehumanization of an out-group can have significant behavioral and attitudinal consequences for political decisions. But what about the dehumanized out-group itself? How does a dehumanized out-group experience being called "vermin" – a term with which for example Polish people living in the U.K. have been targeted with in the aftermath of "Brexit"[49]? And what happens if a majority, high-status group itself feels dehumanized by another out-group? Does the knowledge that one is being dehumanized increase reciprocal dehumanization, and does it affect intergroup conflict and aggression?

Bruneau and his colleagues (2016) pursued this latter question in a follow-up study, where they tested whether 'metahumanization', i.e. to be viewed as less evolved by another group, leads to the outbreak of conflict and violence. "No prior work", they claim, "has examined how individuals respond to the (meta)perception that their group is

[49] http://www.independent.ie/world-news/europe/britain/no-more-polish-vermin-wave-of-hate-crime-and-racial-abuse-reported-after-brexit-34836636.html

dehumanized" (ibid.: 344), even though previous concepts such as 'metastereotyping' exist (Vorauer et al., 1998). Based on the earlier mentioned work of Tajfel, we know that belonging to an in-group and deriving esteem from that membership is a widespread need. Further to this, we know from other studies that negative evaluations of one's in-group can be perceived as a threat that individuals might seek to remedy (Hornsey, Harth, & Barlow, 2011).

Yet Bruneau and colleagues wanted to test specifically whether the explicit dehumanization of another group would affect intergroup relations, and how this would play out in real political conflict situation with large-N samples in a range of cultural contexts, such as the Israeli-Palestinian conflict, the Charlie Hebdo attacks, the hostile relationship between ethnic Hungarians and the Roma minority, and the U.S.-Iranian nuclear deal. For example, Israelis who believed that they were dehumanized by Palestinians (i.e. metadehumanized) were more likely to reject peaceful conflict resolutions and support instead disruptive actions such as population transfer and collective aggression towards Palestinians. In the study, metadehumanization "had significant direct and total effects on all variables, again suggesting its unique role in predicting hostile intergroup attitudes and policies" (ibid., p.355). Similarly, Americans who were told that Muslims viewed Americans as animalistic brutes and less developed than themselves in the wake of the Charlie Hebdo attacks, for example, were much more likely to support torture of Muslims and support of drone strikes as a result.

Finally, Bruneau and colleagues tested their hypothesis with a large-N (906 participants) sample of the Hungarian population in regard to their tense relationship with the Roma people. When ethnic Hungarians were told that the Roma target them for theft

because of the disregard they have for Hungarians' suffering and because they view them as less human than themselves, Hungarian respondents were much less likely to be willing to fund projects for Roma integration, more likely to support discrimination and emotional hostility towards the Roma.

In all of these cases, the knowledge that another group viewed your own in-group as less human and less evolved was a significant contributor to hostile reactions and support for aggressive retaliation. It is important to note here, again, that the feeling of being disliked by a group (i.e. prejudice) was not as significant a predictor of intergroup aggression, nor was subtle dehumanization. This is not to say that either prejudice and subtle dehumanization do not have an affect on political attitudes and behavior, but that blatant dehumanization and the explicitness attached to its rhetoric (i.e. animalistic and degrading terms and explicit hierarchical ordering of different groups) has to be acknowledged as a worrisome and exacerbating factor in intergroup conflict in its own right.

Another interesting aspect of this study is that Bruneau and colleagues tested their metadehumanization theory on high-power, majority groups instead of low-power, minority out-groups. This highlights the need to understand the impact of dehumanization not only on those groups who are usually seen as vulnerable, but also those groups who are usually assumed to be the dehumanizing offenders. It turns out that explicitly telling each other that you do not belong to my sense of humanity and that you believe that the other group is less evolved than one's in-group can actually be politically disastrous for *both* sides. However, if we combine this insight with previous research into how low-power groups care more than high-power groups to be respected and perceived as

competent (Bergsieker, Shelton, & Richeson, 2010), then it seems highly plausible to argue that low-power groups might be especially sensitive and reactive to metadehumanization, and that this in turn can lead to a higher chance of political aggression and resentment.

In the context of the 'civilizationary clash' rhetoric, this becomes significant as we look for ways to establish moral and rhetorical standards of how different groups should talk to and about each other. Not only do we have to pay attention to dehumanizing rhetoric directed towards low-power groups, but we also need to recognize that dehumanizing language and the belief that one's in-group is viewed in a humanly inferior and demeaning way by an out-group (even if that out-group holds less power and status in society) affects high-power groups as well, and can therefore have toxic effects on intergroup peace and cooperation.

I argue therefore that civilizationary clash theories are not only problematic because they might make non-Western nations and cultural groups feel excluded from a shared sense of humanity and humiliated as a result of being portrayed as culturally and economically inferior, but also, because the idea of a hierarchical ordering of different 'civilizations' can backfire for the 'superior' West, in that non-Western groups might feel metadehumanized by this kind of discourse, and as a result, return and retaliate by blatantly dehumanizing the West. The 2016 Bruneau study showed that even high-power groups, i.e. in this case, the West, are very much affected by feeling dehumanized by a low-power group, and that this sense of dehumanization experienced by a high-power group can result in support for aggressive and belligerent actions to resolve conflict, instead of compromise and peaceful negotiations.

One only has to think of the phenomenon of Anti-Americanism in the Arabic world, which is often expressed through flag burning or more recently, violent protests in front of U.S. embassies in the Middle East following the Arab Spring in 2011[50]. Amaney Jamal (2012), in her book on Anti-Americanism in the Arab world, argues that the persistence of Anti-Americanism should not be explained through the existence of some deep civilizationary hatred or as an emotionally charged sentiment towards the world's largest superpower, but as a rational response to U.S. policies and the way how they have systematically disadvantaged Arabs economically and politically. Jamal claims that because the U.S. has insisted on "pro-American democracy or no democracy" at all, authoritarian rule was allowed to flourish in many places, opposition against the latter which Jamal pins down as the main source of Anti-American sentiment. Based on the PBP argument, Jamal's claim could be reconsidered. Indeed, Jamal's attempt at characterizing Arabs as rational decision-makers with sensible political grievances instead of an emotion-driven and brutish mob can be understood as a way to humanize the Arabic people to a American audience.

As much as this kind of re-humanization of Arabs and the reminder that Arabs can be rational actors according to Western standards just as everyone else is much needed in the current U.S. foreign policy discourse, Jamal might be overlooking the relevance and power of civilizationary rhetoric on political behavior. Although Anti-Americanism in the Arabic world should not be pinned down to the existence of some dubious and ominous existence of ancient civilizationary hatred, Jamal might be underestimating how the West's obsession with the civilizationary clash theory can come

[50] http://www.cnn.com/2012/09/11/world/meast/egpyt-us-embassy-protests/.

across as dehumanizing for the Arabic world and in turn, the humiliation and anger in response to both the subtle and blatant dehumanizing by the U.S. towards Arabs might manifest itself partially in the anti-American protests and flag burning events that took place in the wake of the Arab Spring.

In his review of Jamal's argument, Marc Lynch (2013) points out that the role of cognitive bias was overlooked, in that Arabs might ascribe much more powers to Washington in influencing their lives and political events than is actually the case – indeed, with the PBP, this cognitive bias could be explained by the fact that feeling dehumanized and viewed as a lesser and more backward 'civilization' deeply affects group members' sense of worth and can backfire in aggression and retaliation. Ironically, the hated high-power group (in this case the U.S.) would then be ascribed much more cognitive significance and as a result, be viewed as the main culprit for one's political woes.

I contend further in this context that we need to redefine what we mean by 'rational' political responses, since reactions to feeling metahumanized in the form of retaliating by dehumanizing another group in turn could in fact be constructed as rational cognitive mechanism, which serves protective and strategic functions, even though the behavioral and attitudinal outcomes are not politically desirable. From the PBP it is important that we begin to take into account the emotional and cognitive effects dehumanization has on our sense of self-worth, group-membership and humanity, and that once a civilizationary rhetoric about "us" versus "them" is unleashed, dehumanized groups might continue and reciprocate dehumanization in a vicious cycle of political actions and rhetoric. Painting this cycle as exclusively non-rational is just as misleading

as painting it as completely rational (in the way how rational choice theory defines 'rational' agents, for example).

Rather, if we return to a point I made in the introductory chapter and chapter 2 of this dissertation, we need to seriously redefine what 'rationality' means in the intergroup dynamics of mutual denial of humanity, and how everyday dehumanization is part of how our social brains function, but also, how dehumanizing someone and feeling dehumanized in turn is experienced at the brain level in a complex and visceral way that cannot simply be dismissed as 'irrationality'.

A good example of this confusion about 'rationality' is the discourse surrounding the Bosnian war and genocide. The Bosnian conflict took place between two main groups, Bosnian Serbs and Muslims, and resulted in the slaughter and attempted eradication of the Muslim population in the early 1990s. The discourse preceding and surrounding the Bosnian conflict in the form of "Balkanization", i.e. painting the situation as an outbreak of ancient hatreds and irrational feuds that had been going for centuries in this region, was a major contributor in the West's reluctance to intervene (Zimmermann, 1996; Dobbs, 1995; Bush 1992; McCain S1204010, 102nd Congress, August 10, 1992). Two of the most influential publications were Robert D. Kaplan's *Balkan Ghosts* (1993) and Rebecca West's *Black Lamb and Grey Falcon: A Journey through Yugoslavia* (1941). Kaplan's in book in particular was said to have persuaded President Clinton to abandon his policy of 'lift and strike' in May 1993 and adopt the stance that the Balkan conflict was unsolvable due to the civilizationary peculiarity of the region[51]. Lene Hansen (2006) sharply criticizes both Kaplan's and West's works for

[51] Roger Petersen (2011) makes a similar but much more sophisticated and subtle argument about how Western powers have not sufficiently mentalized Balkan people, and that Balkan people's

conveying a 19th century romantic image of the Balkans, which celebrates the Orthodox and Byzantine influences on the Yugoslav region and draws a sharp Western "us" distinction against the Orthodox "them" (ibid., p.153).

Another influential voice in the discourse is the *Carnegie Endowment for International Peace Report on the Balkan Wars*, published in 1914, 1993 and 1996. Unlike the dominant Western discourse that assumed insurmountable civilizationary differences between the West and the Balkan region, the 1914 Carnegie Report took a markedly different position. Although American and Western civilization were declared as superior, the 1914 report struck a very optimistic and humanizing tone, in which Balkan people were described as being "the same as us" and reformable, and thus should be made "young clients of civilization" (in Hansen, 2000, p.8).

Although the 1914 report was written as an attempt at reconstructing history after the Second Balkan War in 1913, and not as a piece to advocate Western intervention, it is important to note the outlook in this report, where the attainment of a singular civilization was seen as possible[52]. In other words, the West's in-group boundaries were defined as flexible and embracing of the Balkan people, and for this reason, the Balkan conflict was not viewed simply as an ancient, unintelligible conflict, but as an intelligible and tragic event carried out by humans just like us. The Balkan people were *not* depicted as animalistic or mechanistic brutes and savages but were instead ascribed characteristics and decision-making capabilities that were intelligible and *potentially equal* to Westerners.

---

motivations and aspirations might not adhere to the highly individualized and rationalist model of Western rational-choice based political models.

[52] See Trix (2014) for a balanced and critical historical analysis of the 1914 Carnegie Report on the Balkan Wars.

On the other hand, the 1993 report was politically much more crucial than the 1914 report, and stood in stark contrast to the latter. Prefaced by George Kennan, it depicted the Balkan conflicts as a savage battle between "emotionally excited" groups whose untamable drives could not be subdued by Western powers but should instead be left to itself to resolve, almost like a wildfire (note the animalistic dehumanization at play here). Kennan's advice not to intervene is decisive in this aspect, urging Western leaders not to get involved in a conflict that is essentially carried out by people with a different kind of (i.e. more brutish) humanity than the West's. In comparison, the final 1996 report, written up in the aftermath of the war and after the shocking revelations of genocidal violence directed against Bosnian Muslims, takes a more balanced stance by arguing on the one hand that the West could have and should have intervened – criticizing in particular the UN's portrayal of the war as a natural disaster – yet on the other hand being reluctant to revert to the 1914 report's optimistic belief in the progress of universal civilization.

Critics of the 'ancient hatred' theory – some of whom were actual witnesses to the Bosnian conflict – strongly condemn the civilizationary perspective for denying the actors in the conflict rationality and intelligible political motives. For John Mueller (2014), for example, the Bosnian war was not a manifestation of internal civilizationary clashing within the region but a political event that happened as a "result of inadequate government" (ibid., p.47). Similarly, Hansen contends that "the war should be understood in the terms of Serbian nationalist aggression directed against a tolerant and liberal Bosnian government and its citizens" (ibid.: 346). Hansen criticizes harshly the construction of the specter of "Balkanization", which portrays the region as hopelessly

embroiled in intergroup violence and its people as "prey to (…) violent promptings of their own passions", which would entrap the Western great powers if they were so foolish as to intervene (Todorova, 1997: 34).

Ed Vuillamy (1998), one of the major British journalists who reported about the Bosnian conflict from the ground, offers a more nuanced explanation of the conflict's causes. Vuillamy believes that one of the main reasons why Serbs felt threatened is because they were told by their political leaders that Muslims were trying to kill them and eradicate them, since "the first task facing any group intending to inflict genocide on someone else is to convince its own people that they are about to be victims of genocide themselves" – indeed, the 'genocide' term was first used in the context of Serbs fearing expulsion by President Milosevic (ibid., p.77). In addition, Radovan Karadzic, the main political instigator and war criminal on the Serbian side, was a rural *papak* who felt rejected and betrayed by a cosmopolitan Sarajevo dominated by urbanized Muslims, and possibly even felt dehumanized by that very community. In other words, Vuillamy's observations suggest that Serbs might have felt *metadehumanization* prior to their own instigation of genocide of the Muslims, which in turn led to the perceived need for retaliatory and excessive violence against the Muslim out-group.

Instead of being driven by ancient and opaque hatreds, this shows that Serbs might have been reacting aggressively and violently to Muslims because they had worked themselves up into believing that they were dehumanized victims. They were fuelled by the fear-driven *metadehumanizing* rhetoric of their political leaders, to the point of feeling sufficiently dehumanized and threatened themselves that they could justify retaliation against the Muslim out-group.

The PBP on dehumanization is crucial in the Bosnian context because it firstly debunks the civilizationary divide theory by taking on the perspective that we all share the same brains cross-culturally, therefore political scientists cannot dismiss certain parts of the world and its people as cognitively unintelligible. Instead, the PBP compels political scientists to come up with explanatory models premised on the humanly intelligible motives of its political actors. Secondly, the psychological and neuroscientific research into intergroup violence and dehumanization offers a strong glimmer of hope that dehumanization can be tackled through various mentalizing and empathy efforts, as well as cooperational set-ups, which in turn makes the Bosnian conflict seem less like an endless vicious cycle of violence but a toxic intergroup conflict that could have stood a chance of de-escalation and intervention, if only political leaders and experts would have taken seriously the devastating political consequences of feeling dehumanized by another out-groups.

5.3.3. Why the PBP on Civilization Theories Matters III: International Human Rights

Civilizationary clash theories based on hierarchical distinctions between "us" and "them" do not just appear in U.S. foreign policy debates and the International relations field but also in human rights theory. I contend that International human rights scholars often operate on a similar logic in order to justify the implementation of human rights and Western norms abroad. For example, a widely respected authority on human rights theory, Jack Donnelly (1998), couches his explanation of the growing acceptance of universal human rights in the post-War period (and post-Cold War period in particular) in the language of divisive civilizations. Human rights, he says, emerged as a new

international "classic standard of civilization", in which being inside or outside is defined through the acceptance of shared cultural values.

He traces back this civilizationary standard exclusively through Western history, claiming that this standard first emerged in the 19th century in the context of colonialism and the question of extraterritoriality, in which imperialist countries such as Britain needed to find a way to distinguish between "uncivilized natives" in prospective colonies who could then be justifiably civilized through colonization and excluded from international law, and on the other hand, countries such as Imperial China, which could clearly not be described as uncivilized or savage but was also not allowed membership in the "family of nations" headed by the West (ibid.: 4). The latter case was understood to fall under the realm of "extraterritoriality", i.e. applying to countries such as Imperialist China, Japan and the Ottoman Empire, which were not considered to be at the same "savage" level as the African dark continent, but at the same time, were not accepted as being as civilized as the colonialist West and hence treated as sovereign, but still unequal. It is important to note here both the subtle and blatant ways in which "uncivilized" and "extraterritorial" out-groups were being dehumanized in the 19th century.

Africa and other "uncivilized" regions of the world were blatantly dehumanized by being portrayed as brutish, savage and ape-like, whereas "extraterritorial" cases such as China or the Ottoman Empire – even though they were deemed sovereign – were subtly dehumanized or infra-humanized by denying them uniquely human emotions and attributes through "Orientalism" and being given the status of exotic exceptionalism (Said, 1978). To put it more bluntly, the world regions that the Imperialist West colonized were simply written off as beneath their own humanity, whereas other world

regions that the West did not dare to aggressively colonize were characterized as exotic freaks that might have made certain cultural, economic and political achievements, but were lacking in a special human essence that only the West possessed[53].

There are parallels to this kind of subtle dehumanization of the East with the reported mechanistic dehumanization experienced by East Asians in the West today (Haslam, 2006), which was brought up in Chapter 3 and 4 already. East Asians who are mechanistically dehumanized are seen to lack unique human traits such as emotional warmth, compassion, individuality and creativity and are often perceived (and feared) as over-achieving, excessively competent, high-functioning machines. Subtle dehumanization of this particular group is therefore repeating itself in a strikingly similar fashion to colonialist periods in Western history – casting doubt on whether a cognitive rejection of dehumanizing perceptions of East Asians ever took place.

Despite the West's troublesome colonialist usage of the idea of civilization and hierarchical civilizationary divides between different groups, Donnelly presses on in his advocacy of modern-day human rights as a desired "standard of civilization", arguing that in the 20th century, the West was able to liberate itself from the ugly shackles of colonialism and turning towards a more egalitarian idea of sovereignty. Donnelly

[53] This difference can also be analyzed at through the lens of "tutelary politics", in which case the "other" is infantilized rather than dehumanized. For example, J.S. Mill, who worked for the East India Company, argued that Indians were not ready for self-government, although they could be ready for self-governance eventually (Zastoupil, 1994). Mill famously distinguished Indians from "savages" (the lowest rank within the civilizationary scale), classifying them instead as either "semi-barbarous" or "barbarous" (Tunick, 2006). This begs the question whether instead of having blatantly dehumanized Indians, Mill held an infantilized view of the Indian people who however to him still had some level of agency in determining their history and changing their government at some point in that history. The ascription of human agency and the capacity for bringing about historical change thus seems to be an important factor in evading more extreme forms of blatant dehumanization and categorization of people on the lowest rank of the civilizationary scale.

contends that the "standard of civilization" was suddenly turned into a force of good, such as igniting campaigns against the slave trade, advocating for penal reform, outlawing practices such as piracy, polygamy and infanticide (ibid.: 5). Historical milestones, such as the when the League of Nations pledged after WWI to protect the rights of national and religious minorities in their territories and most important, the adoption of the Universal Declaration of Human Rights in 1948 by the United Nations General Assembly, serve as legal and normative examples of this "standard of civilizations" for Donnelly. Based on these achievements, as well as the success of the colonial independence movements in the latter half of the 20th century, Donnelly perhaps too optimistically (and naively) concludes that "the whole globe was recognized as civilized" in a postcolonial era (ibid., p.13).

This chapter casts doubt on this optimistic conclusion, not least based on Bruneau and his colleagues' (2015; 2016) studies discussed earlier, which contradict Donnelly's claim completely by showing that civilizationary concepts based on dehumanizing and divisive hierarchies very similar to those of the darker colonialist periods in Western history are still alive and potent within the West today. The most troubling part of Donnelly's argument, however, lies in the part when he insists that we need to continue to uphold the "standard of civilization" because "something like a standard of civilization is needed to *save us from the barbarism* of a pristine sovereignty that would consign countless millions of individuals and entire peoples to international neglect [emphasis added]" (ibid.: 16). He invokes the recent tragedies of Rwanda, Bosnia and Tiananmen Square, where suffering populations were helped with too little, too late by the international community and were left at the mercy of their national and local politicians

to die unspeakable deaths. Donnelly is worried, just as other interventionist human rights advocates (Power, 2002), that the abandonment of a "standard of civilization" – both as a rhetoric tool and a normative international measure – will be tantamount to conceding defeat to human rights abuses and atrocities committed by authoritarian and oppressive regimes. Donnelly ascribes to the concept of "civilization" and the drawing of the dichotomy between the "civilized" and the "barbaric" a great persuasive quality – in the hope that the mere invocation of the term will spur the international community into action.

I claim that there exists a strong connection between Western political theorists defending of an essence-based idea of dignity on the one hand (as discussed in Chapter 4) and of this kind of civilizationary argument espoused by Donnelly. The hope is that by simply stating and believing that a "standard of civilization" and "universal dignity" must exist, international human rights laws and interventionist actions taken against oppressive regimes will have the necessary fuel and justification to be carried out and survive. My argument is that Donnelly, by invoking a "standard of civilization" against the "barbaric" other, might in fact be engaging in potentially subtle and blatant dehumanizing of those very out-groups he wishes to help, and therefore, in a counter-productive fashion, invite people to engage in dehumanizing perceptions of various groups and countries in the world.

Although Donnelly's intentions are opposite to those of Huntington, in that Huntington is wary of interventionist politics in non-Western countries because he fears a battle between incommensurable cultural values (whereas Donnelly advocates for interventionist politics in the name of universal human rights), both Donnelly and

Huntington *completely misunderstand what the invocation of the "civilizationary divide" does to the social brains* of both Westerners and non-Westerners. Instead of making people more wary of interventionist politics in non-Western regions (Huntington's aim) or more empathetic and proactive in intervening in human rights violations abroad (Donnelly's aim), the studies on subtle and blatant dehumanization unmistakably show that individuals who dehumanize someone through the civilizationary lens are less willing to help that out-group and are more likely to engage in aggressive actions. In addition to this, individuals who feel that their own in-group has been dehumanized want to retaliate belligerently against the other group and refuse cooperation with them.

In a nutshell, the outcome of thinking of people as less evolved and barbaric, i.e. as less than human than ourselves, and of thinking of ourselves as superior on the civilizationary ladder *makes us all worse off politically*: those who dehumanize others as a result lose the capacity to mentalize others through deactivation of their medial prefrontal cortex and exhibit negative and uncooperative political attitudes and behavior, and those who feel dehumanized might then dehumanize the other group in return and retaliate violently.

It is this *retaliatory dehumanization* by the initially dehumanized group that is often just marginally discussed in research on intergroup conflict, particularly how it plays out within the social cognition systems of both dehumanizer *and* the dehumanized (Ascher & Mirovitskaya, 2015; Dancygier, 2010; McDoom, 2012; but then see Claassen, 2016). Although it is important to foremost outline power and status imbalances that determine dehumanized perceptions of certain out-groups by more powerful and higher-status groups (Harris & Fiske, 2008), it is also necessary to acknowledge that

dehumanized groups can dehumanize their dehumanizers in turn – in other words, the politics of dehumanization can be a two-way street, even if it is often a politically and socially unequal one.

Donnelly is therefore terribly naïve in contending that the dehumanizing and undesirable aspects of the "civilization" concept back in the 19th colonialist century can be easily overcome in the 21st century, by claiming that the idea of a superior "standard of civilization" does not always have to be correlated with power and domination (ibid.: 5) and that we have "moved a significant distance from civilizationary imperialism" (ibid: 21). Critics of Donnelly within the IR field reject the claim that we have entered a postcolonial and postimperial age but believe that many contemporary practices, including within the IR field itself, are driven by economic and political neocolonialism (Jones, 2006; Hall and Jackson, 2007). Julian Saurin (2006) argues that if we, according to Lenin, understand colonialism as chiefly an expression of capitalist imperialism, then in many ways, colonialism did not come to an end with the national independence movements of various former colonies (ibid.: 26). Instead, Western supremacism continues in the structural production of an economic world order skewed in favor of the West and the insistence on the nationalist principle as the primary way to gain political legitimacy on the historical stage. Saurin concludes that from the moment of its conception, the field of IR was imperialist and national in character, founded on "myths of nationalist origination" and the aim to monopolize the means of mental production around a Western vantage point of superiority.

Branwen Gruffydd Jones (2006) goes so far as to say that IR studies are a form of "modern imperial ideology" (ibid.: 5) in which international relations are predominantly

conceived as imperial relations. She uses the example of Islam as a way how Western IR studies have both misunderstood and caricatured the phenomenon in its religious, historical and socio-political complexity, and how there is a "deeply rooted, almost subconscious tendency to deny legitimacy and worth of non-Western values, traditions, practices, struggles, discourses, and thought" (ibid.: 12). The biggest criticism for Jones, which could be applied to universal human rights theorists such as Donnelly, is that IR is so self-confident about its own good intentions towards non-Western countries and groups that it comes to believe its own "myths" and the construction of a "privileged, genealogically useful past, a past in which we exclude unwanted elements, vestiges, narratives" (Said, 1994: 16).

This chapter treats the neocolonialist critique by IR theorists on the Left as complementary to the previous neuropolitical criticism of Donnelly. The neocolonialist critique within IR chiefly uses materialist and historical arguments to make their point, which provides a crucial backdrop to the power relations between the civilized "us" versus the barbaric and backward "thems" inherent in the political civilizationary clash theories, as well as the dehumanizing civilizationary hierarchies detected at the brain-level. Meanwhile, it could be argued that the "Ascent of Man" ratings in the Bruneau et al. study followed a pattern of global economic status hierarchy, which might be correlated with the kind of Western supremacist and neoimperialist bias highlighted by its critics.

Donnelly's stance is part of a larger call within the human rights debate to "name and shame" human rights abusers, if necessary with civilizationary arguments (Farer, 2000; Franklin, 2008; Krain, 2012). Indeed, what bigger shame and loss of face could the

international community bring to another country's leaders by calling them "savage", "barbarian" and denying them sophistication and humanity? It seems to be a recurring phenomenon that dictators throughout history have aspired to be popular and loved, which might however be more due to individual egomaniac vanity than to having read and taken to heart Machiavelli's advice to princes that they should always try to be loved instead of being hated. More seriously, the idea here is that by shaming internationally a government or political actors who commit human rights abuses against their citizens, the citizens' own perception of their governments' crimes will shift, thus increasing the likelihood of rebellion and dissent (Gurr, 1970; Lichbach, 1995).

The empirical evidence on whether "naming and shaming" is actually effective is mixed, with some studies showing that this tactic can lead to a significant shift in public opinion (Ausderan, 2014) and others suggesting that shaming can in fact result in an overall increase of violations (Hafner-Burton, 2008; Krain 2012). I argue here that shaming has to be done very carefully, with a special sensitivity towards dehumanizing language. In the case of shaming human rights abuses committed by the Chinese government, for example, Western powers and human rights organizations constantly have to manage a balancing act where the "naming and shaming" directed against the government does not trigger an indignant nationalist response from the populace (Hughes, 2006; Liu, 2007).

This was especially relevant in the run-up to the 2008 Beijing Olympics, where international pressure to improve China's human rights record drove some parts of an offended population into rallying behind the government (Yang, 2008). Ongoing hot button issues such as Tibetan and Uyghur independence (and in connection to this,

furious nationalist responses by both young Chinese at home and abroad) suggest that the West needs to pay attention in which way it delivers its human rights critique. Based on the PBP of both in-group and out-group, I wish to stress the importance of making the criticized out-group feel fully humanized and included in the critic's sense of humanity. Both subtle and blatant dehumanization have to be avoided at all costs, since this would trigger defensive and even retaliatory reaction from a Chinese populace's side, which might in fact not be strongly supportive of its government, but feel driven into supporting it as a result of feeling dehumanized by the West.

In this sense, I wish to transcend the current divide within the human rights field between those who believe that human rights standards are universally valid and applicable (Donnelly, 1998; 2007), and the growing oppositional voices that insist that non-Western societies have their own, home-grown human rights traditions and values (Midlarsky, 1998; Sen, 1997; Bauer & Bell, 1999) and that claims to the universal applicability of human rights is deeply Eurocentric (Mutua, 2016; Nisbett, 2003; Zakaria and Lee, 1994). I believe that this divide might never be solved, in that both sides hold a deep historical and political investment in their cases, which can be equally persuasive depending one one's own political agenda and outlook on dialectic historical processes.

What can be said concludingly however, is that the universalist camp, if it wants to succeed, needs to understand the danger of invoking civilizationary arguments, especially when they are supposed to serve "shaming" purposes. Even if the universalist camp disagrees with those who accuse it of Eurocentrism and lack of understanding of local, non-Western traditions, it cannot revert to a hierarchical idea of civilizationary

world order without risking rejection, offense and potentially reciprocated dehumanization.

5.4. Conclusion

The argument I tried to make in this chapter is both simple and complicated. The complicated part consists in the many facets in which dehumanization can manifest itself – either as animalistic or mechanistic dehumanization, subtle infrahumanization, implicit bias or blatant dehumanization. Researchers are still in the process of parsing out the differences between these modes and exploring to which extent they affect political attitudes and behavior. It is for this reason, as well as the particularity in which intergroup dynamics play themselves out in specific historical and situational contexts, that the exact causal relationship and correlation between civilizationary theories and intergroup aggression is still in the process of being developed on various interdisciplinary fronts in the cognitive and social sciences. However, what I wish to have transpired in this chapter is a simple message, which is that civilizationary theories have to be employed with a lot of caution by politicians and political scientists alike. This is because these theories tap into and amplify our natural tendency for in-group favoritism and potentially, dehumanized perception of out-groups. In the end, framing one's political worldview in this way makes both the dehumanized and the dehumanizer politically worse off.

As many non-Western parts of the world are in the middle of a precarious balancing act between modernization and the struggle to define their identities on a global stage, we need to be cautious with the concepts and language employed in international conflicts and interventions. The danger of the civilizationary rhetoric is its

deep divisiveness along the lines of those who are considered to be fully human and those who are seen as barbaric and inferior to another group's superior humanity. We need to remind ourselves that the wish to be humanized by others is a need and passion that runs so deep that it can both heal seemingly insurmountable divides but also become the toxic fuel for a spectacular kind of redemptive revenge. The goal of an identity politics in the 21st century ought to be to strive for the former and avert the latter at any cost.

6. Conclusion

6.1. The Two-Realities Problem

The wider purpose of this dissertation was to show that the brain matters for politics. This is not only due to the fact that the brain is located as a vital organ inside the body of political actors, but also, because certain social brain abilities and mechanisms are fundamental to the dynamics underlying political cooperation, especially in hyperdiverse societies. Yet political theories on diversity and difference commonly suffer from the two-realities-problem: there is one reality imagined by political theorists, based on analytic and normative assumptions derived predominantly from Western political thought, on how individuals from different ethnic, cultural and religious groups can live and cooperate together; and a second reality, in which the actual cognitive and affective experiences of the perceivers of difference and those who are perceived to be different are largely unknown.

The two-realities-problem poses a genuine problem for political theorists: theories about cooperation between diverse groups are always based on some assumption, however minimal, of how human beings operate in social settings with others – yet what if our theoretical assumptions do not match up with the second reality of the social brain? What if, as a result, this dissonance leads to the depiction of a reality that is far removed from how the excluder herself experiences acts of exclusion at the brain level? Further, what if normative demands made by political theorists on how we should include others might be cognitively impossible to achieve, due to our 'flexible social cognition' (Harris,

2017) regarding the everyday dehumanizing of others? Finally, what if political theorists' largely monocultural concepts of self and human dignity are challenged by recent insights from cultural neuroscience on how bicultural individuals switch between both Western and non-Western identities? How should new models of the hyperdiverse political self be structured accordingly?

In an effort to bridge the chasm between these two realities, this dissertation tried to develop the PBP as a new epistemological tool to navigate between insights from psychology and brain data on the one hand, and political levels of analysis on the other. In a post-Civil Rights age where political and social norms about prohibiting the explicit expression and endorsement of racism, sexism and other kinds of prejudices against minority groups are largely accepted and adhered to, the PBP on the implicit exclusion and subtle dehumanization of other groups can become a crucial window into forms of less explicit and yet still persisting systematic discriminatory and exclusionary practices.

Indeed, in an effort to appeal to a larger voter base and to submit to these norms, even current leaders of far-right political parties and movements, such as and Marine Le Pen in France and Steven Bannon in the U.S., do not want to be branded as "racists" and claim that they do not hold racist political views (Polakow-Suransky, 2016; Wolff, 2016). In a reverse example to this, Chapter 3 discussed the fact that liberal thinkers such as Immanuel Kant were willing to reject slavery on ideological grounds but nonetheless viewed Blacks as humanely inferior to European Whites.

It is this split between the content of political ideologies and the way that our brains (do not) enact them that this dissertation tried to draw attention to. If an explicitly unequal and subhuman treatment of any particular group is prohibited ideologically in

one society – and if, as scholars in the post-Cold War era have argued (Fukuyama, 1992), liberalism's normative prescription on social cooperation and diversity is the main survivor in today's ideological landscape – then having access to the second reality of how these norms and ideological beliefs are in fact instantiated in the brains of individual political actors becomes indispensable in gaining a complete picture of political reality itself. The window into the second reality allows us to understand the nature of political beliefs and ideologies beyond their mere written and rhetorical manifestations.

On a more pragmatic level, being able to offer evidence of this second reality can be empowering for minorities and targets of exclusion in instances where, based on the first reality, they are told that exclusion does not exist or that they are just imagining the exclusion that they experience subjectively. The brain evidence is a powerful tool to employ in the context of trying to understand and outline what is happening in the brain of the excluder, and to help establish a reality that minorities experience at an intense, visceral and damaging level, yet whose very existence they often struggle to describe and validate.

6.2. Towards a Post-Ontological Idea of the Political Self

A central assumption underlying the PBP is that in order to understand and predict political behavior, and the cognitive conditions under which individuals enter political scenarios such as the social contract or cooperation in hyperdiverse societies, a minimal neuropolitical theory based on the PBP needs to commit to a non-ontological, non-essence based idea of the political self. I outlined this based on more general epistemological grounds in Chapter 2 and by applying this stance more concretely to the

concept of 'human dignity' in Chapter 4: instead of treating the nature of political virtues and moral goods such as dignity, toleration and recognition as an ontological essence that resides within each of us like an invisible nugget, this dissertation argues that we should instead construct the existence of dignity as an *interpersonal, neurocognitive ascriptive process* that is realized through flexible brain mechanisms of ascribing or denying humanness to someone else.

One fear amongst critics of the mechanism-based approach (Bennett & Hacker, 2003; Nagel, 1974) is that constructing an idea of self based on our 'social cognition network' and the neurocognitive mechanisms that underpin it will lead to extreme reductionist, determinist and non-dualist views of human subjective experience and irreducible goods such as dignity. This dissertation tried to show, primarily in Chapter 2's historical depiction of the highly thoughtful internal discourse within social and political neuroscience, that neuroscientists themselves are usually very skeptical about extreme reductionist and determinist explanatory models of social behavior, and furthermore would balk at claims that brain data alone could explain complex social or political phenomena in their entirety. However, what a mechanism-based model of the political self offers us – especially on topics such as discrimination, prejudice and exclusion – is the ability to access implicit biases, ambiguous stereotyping, as well the spontaneous and rapid way we deny humanness to others through crucial brain mechanisms that would otherwise go undetected through more behavior or survey-based methods.

Social neuroscientists today might be willing to affirm that the 'biological is social and the social is biological', however, this dissertation expressed uncertainty over the question whether the same goes for biology's relationship to politics. Although the

political is certainly always biological, in that every political actor possesses a brain and body (and enacts politics through them), it is not clear whether the biological is also always political. Chapter 3 made the argument that public representatives in a liberal democracy have a special neuropolitical responsibility to mentalize their diverse constituencies but that this neuropolitical responsibility cannot be justifiably demanded from every citizen to same degree. In other words, whereas the existence of unchecked implicit biases and dehumanization of another by a public representative has serious political consequences for targeted groups in terms of power and economic equality, private citizens might hold implicit biases that should ideally be overcome, but that the state or other citizens cannot legally or politically force another private citizen to rescind.

This opens up the question – touched upon throughout all chapters – on what the relationship is between rapid and implicit biases and dehumanizing tendencies, and explicit political behavior. I argued that implicit biases and subtle dehumanization matters for the healthy fabric of liberal democracies and the neuropolitical capabilities of both its members and public representatives. However, solving this question is beyond the scope of this dissertation as a whole, but it is worth mentioning in this concluding chapter that addressing the problems and confusions underlying the relationship between implicit brain mechanisms and actual behavior – and especially its political implications – is an important future research question for the nascent field of political neuroscience and neuropolitical theorizing. Constructing a post-ontological, brain mechanism-based idea of the political self is a first step towards this aim.

6.3. Humanizing the Other: A Basic Condition for Cooperative Politics

The core political argument of this dissertation was centered on exposing dehumanized perception as a major disrupter of cooperative politics in hyperdiverse societies. The neuropolitical reason for this is that dehumanized perception disables a core social brain function called mentalizing, which plays a crucial role in executive and moral decision-making, perspective taking, and feeling empathy with others. To put it more bluntly, if we dehumanize someone in the form of de-mentalizing them at the brain level, they are in our eyes equal to a non-human object (for example, a chair) – with consequences for our ability to fathom their interests, feelings, sufferings and aspirations, as well affecting our behavioral response towards them. Chapter 3 therefore argues that the ability to mentalize a wide array of viewpoints and identities is of utmost importance for public representatives in liberal democracies, whose positions are based on the principle that they will represent their diverse constituencies, as well as to help guarantee equal and fair access to legal and political institutions to all political members of their community. This dissertation argues that without being informed and in command of their 'flexible' dehumanization abilities, public representatives might struggle to fulfill their duties at the cognitive level, and their constituencies might suffer as a result.

In addition to this, this dissertation believes that dehumanization needs to be recognized as a potential disruptor of political cooperation and solidarity within political theories about the social contract, multiculturalism and the politics of difference. I argue that without understanding this fundamental way of how our social brains navigate themselves on a daily basis with others, any political theory about political cooperation situated in the context of hyperdiverse societies lacks explanatory and predictive power. The point is to realize that even well meaning individuals who affirm liberal and

democratic political values can succumb to everyday dehumanized perception of others in rapid and spontaneous ways. It is important to realize that humanizing someone is not an essence-based value that we can pledge allegiance to at one point in our lives but that overcoming our inherent brain-based ability to dehumanize others has to be a conscious, daily effort – being in command of our 'flexible social cognition' thus becomes a chief neuropolitical condition for successful cooperation in hyperdiverse societies.

However, even though dehumanized perception and biases are such a fundamental part of our social cognition system and cannot be completely overcome, it can be moderated and kept in check through a mundane cognitive exercise such as mentalizing other people's food preferences (Harris & Fiske, 2007), exposure to counterstereotypical exemplars (Lai et al., 2014), forming more individuating impressions of others (Fiske, Lin, & Neuberg, 1999), stressing what we have in common on a human level, and mindfulness training of one's empathic and kind feelings towards others (Fredrickson et al., 2008). Activating our reward and competition cognition network can also help towards humanizing others and treating them as people who deserve our concern, such as by shifting group boundaries through competition (Lai et. al, 2014) and establishing a sense of "cooperative interdependence" between groups (Ames & Fiske, 2013). Understanding the cognitive mechanisms underlying dehumanizing is however the necessary first step in working towards trying to partly overcome it. Humanizing others cannot be taken for granted or treated as a by-product of liberal democratic values such as toleration, recognition and inclusion, but instead needs to be treated by political theorists of diversity and difference as a *cognitive challenge for the political self* and as a fundamental neuropolitical ability that needs to be in place in order

for the above mentioned liberal democratic values to have a chance of realization and long-term survival in the first place.

Whereas Chapter 3 and 4 focused on the need for humanization and mentalizing of others in domestic contexts within liberal democracies, Chapter 5 situated this argument in the international political arena, looking at the post-Cold War discourse on a 'civilizationary clash' between Western 'civilized' nations and non-Western 'barbaric' ones, as well as the way how groups who feel dehumanized by others in international conflicts have the potential to engage in *retaliatory dehumanization* in return. The point there was to underscore the importance of the PBP on dehumanized perception not only for the politics of domestic contexts but also for international conflicts and the language used in diplomacy between Western and non-Western states.

Most important, Chapter 5 exposed the disastrous behavioral outcomes of feeling dehumanized – from supporting retributive justice to engaging in intergroup violence – and the way how ascribing a unique human essence to one's national (or Western) in-group in the form of infrahumanization can seriously hamper efforts for effective international dialogue and cooperation to take place. In the case of the international political arena, words and rhetoric surrounding 'civilized' in-groups and 'barbarous' out-groups truly matter because they have an effect on both the brain of the speaker and brains of individuals and groups who are addressed and portrayed as 'barbarous'. Recent statements by U.S. military leaders, such as current National Security Adviser Lt. Gen. H.R. McMaster on how using the term "Radical Islamic Terrorism" is counterproductive

because of how it associates terrorism with Islam as a religious group seems to support this argument[54].

Being able to mentalize and humanize other groups and nations on the international stage, as well as feeling humanized by others internationally is a necessity and fundamental brain-based need that IR scholars and analysts have to take into account when trying to understand what motivates the outbreak of international conflict. Whereas dehumanization can destroy the fabric of a liberal democracy at the domestic level, it can lead to potentially avoidable escalations of retributive and retaliatory violence between different cultural groups and nations. Stressing what we have in common in terms of our humanness, whilst still acknowledging differences in terms of political and economic interests, seems crucial in avoiding the dangerous effects of mutual dehumanization on the international stage. In the context of another international crisis – global warming and environmentalism – this approach of affirming shared humanness on a global scale can possibly even contribute to a higher willingness to proactively engage in the fight against the environmental degradation of the globe (Kashima & Margetts, 2014).

6.4. Neuropolitics of the Future: Cognitive Warfare versus Cognitive Solidarity

In a recent investigative journalistic piece for *The Guardian* newspaper, Carole Cadwalladr[55] exposed how one of the main donors to Donald Trump's 2016 presidential campaign – data science and hedge fund billionaire Robert Mercer – employed a sophisticated internet-and social media based system during the 2016 presidential

[54] http://www.cnn.com/2017/02/25/politics/nsa-radical-islamic-terror-term-unhelpful/
[55] https://www.theguardian.com/politics/2017/feb/26/robert-mercer-breitbart-war-on-media-steve-bannon-donald-trump-nigel-farage.

election, which collected and analyzed social media and Facebook data based on "bio-psycho-social profiling", and targeted profiled users with the intent to influence their political views at an emotional and cognitive level. Also described by its creators as "cognitive warfare", the strategy is to combine big data analysis, cognitive science and new social media technology to understand what motivates voters from a psychological and brain-based level (and how, based on this information, they can be influenced politically through social media messaging). Although using profiling and targeted messages to reach voters is certainly not a novel phenomenon in the history of modern politics (Hillygus & Shields, 2008), the *Oxford's Internet Unit for Computational Propaganda Project* deems the concept of 'cognitive warfare' in conjunction with big data and social media to be a significant development in "computer-scripted computational propaganda", where biocognitive profiling and social media bots are used to influence public opinion[56].

This further underscores the relevance of the PBP and the urgent need for political science to include neuroscientific insights into our understanding of the political brain. Political scientists can no longer choose to ignore the ways in which brain insights and politics are linked in 21st century election campaigns, where, due to increased social media and internet usage by large parts of the voting population in the developing world, as well as the availability of increasingly personal and private 'biosocial' information

[56] For more information, see their webpage at: http://politicalbots.org/?page_id=24. See also http://www.aljazeera.com/indepth/opinion/2017/02/russia-soft-warfare-cyberwar-hackers-fake-news-170227070148722.html. For a more scholarly treatment of "extra-factual" sources of political information (EFI) and states' decision-making, see Kelly Greenhill's current work. For a recent assessment of the cognitive abilities and limitations of voters, and the way how social identities influence voting decisions, see Achen & Bartels (2016).

during this usage, unprecedentedly large and extensive data sets on people's psychopolitical proclivities are collected. Without a general understanding of the cognitive mechanisms behind voters' decision-making processes and a special awareness of the role that dehumanized perception plays in the formation of partisan in-group allegiances and out-group disgust towards groups such as Muslims, immigrants and refugees, political scientists will struggle to make sense of developments at the nexus of big data, cognitive science and internet technology that are poised to define the political landscape for this century.

In opposition to the idea of 'cognitive warfare', I wish to close by sketching a vision for what I call *cognitive solidarity*. As we have seen, the picture that has emerged of our social brains does not necessarily fit very well with the normative political demands of a liberal democratic society: our brains are prone to in-group favoritism over equal and neutral treatment of all; we hold inherent biases and prejudices towards other out-groups; we succumb to heuristic biases when trying to make rational decisions; and we spontaneously and rapidly ascribe or deny humanness to others, even though we might consciously commit to universal and cosmopolitan values.

Despite the fact that the initial picture of our social brains appears to be rather dark and sobering, this does not automatically mean that our brains are unable to function well and cooperate with others in a liberal democratic setting. In Chapters 3 to 5, as well as earlier in this conclusionary chapter, I outlined the various ways in which we can challenge our brains to overcome biases and dehumanized perception, and also pointed to the plastic and malleable nature of our brain structures and function. This dissertation's main aim was to draw attention to dehumanization as a serious threat to political

cooperation in hyperdiverse societies, and the obvious next step would be to investigate ways to overcome this threat, at least to some extent.

I argue in these final paragraphs that before dehumanization can be overcome politically, we need to forge a new neuropolitical concept – such as 'cognitive solidarity' – in which the cognitive challenges to mutual humanization and inclusion are acknowledged, but most importantly, in which the idea of cross-cutting solidarity between different socio-economic, racial, cultural and to an extent, political groups is treated a highly precarious yet acutely needed political good.

A neuropolitical concept of 'solidarity', I argue, is distinct from simply asking people to be more empathetic towards others or overcoming their dehumanization tendencies toward certain groups. Solidarity here is understood to be more than a (potentially fleeting) emotional state or an individualist way of taking control over exclusionary aspects of our 'flexible social cognition': rather, it expresses a distinct political commitment in that it contains a political justification over why someone ought to extend inclusion and their cognitive resources towards another out-group. It addresses the question of "What's in it for me?" in that it tries to make a persuasive argument about the merits of cross-cutting solidarity for individuals and the particular groups they belong to, and offer incentives to pursue solidarity as a worthy neuropolitical good for all members of society.

Social scientists and historians are know to criticize neuroscientific and psychological concepts of social emotions and behavior for lacking historical context, external validity and direct political relevance (Bauman, McGraw, Warren, & Bartels, 2014; Sered, 2014). In the introductory chapter of this dissertation, I argued that the brain

data alone is not politically self-evident – this also applies to the psychological and brain-based evidence on overcoming exclusionary biases: just because bias reduction works well in experimental laboratory settings through playing cooperative games does not mean that simulating the same kind of game in the real political world will yield similar results in the long term; likewise, just because empathy seems like a desired emotion to foster between individuals and groups does not automatically make empathy itself a desired and persuasive political good.

This is why a PBP on 'cognitive solidarity' would be different than just asking people to mentalize, humanize or empathize with each other: the idea of solidarity is distinctly political and not just attitudinal or emotion-based – in that solidarity appeals not only to our brains in their individual biological existence, but situates the fate of our brains (and the human beings attached to them) in a historical and political community of other fellow individuals who, just like us, cannot accomplish their political objectives on their own. Affirming the need for 'cognitive solidarity' in a neuropolitical fashion means that we acknowledge that we need to overcome certain brain-based biases and dehumanization tendencies in order to achieve our own and our group-based political goals, and that humanizing other out-groups is a first step towards it. This incentive-based, long-term and communal politico-historical outlook on overcoming exclusionary biases is supported by other neuroscience research in moral decision-making, for example, which shows that basic reward-based neuronal networks underpin moral decision-making processes in individuals (Greene, 2015).

From within political science, the topic of solidarity and collective action has attracted long-standing attention. In particular, Seymour M. Lipset's *Political Man*

(1960) stands out in its analysis of the importance of "cross-cutting cleavages" for the foundations and success of democracies around the world, which was subsequently validated empirically a couple of decades later (Mutz, 2001; Selway, 2011). This further supports the need to investigate and flesh out the idea of 'cognitive solidarity' for hyperdiverse and divided democracies of the 21st century. Approaching solidarity from a political perspective also allows us to questions whether an 'individuation' based approach to overcoming dehumanization (Swencionis & Fiske, 2014), i.e. humanizing someone by perceiving them as uniquely individual is sufficient in maintaining exclusionary biases towards them in long-term political and institutional contexts. I argue that we need instead an incentive-based argument of 'cognitive solidarity', which should not merely rely on cognitive individuation strategies but also on distinctly political and communal incentives for people to join. This discussion goes to show that the PBP developed in this dissertation is useful not just for understanding exclusion and dehumanization, but also, for devising neuropolitical theories for other relevant problems in liberal democratic politics.

In the introductory chapter, I quoted Thomas Kuhn for stating that one of the clearest signs that a true paradigm shift took place within a discipline is that it changes the problems that become available for scientific investigation and reflection. Kuhn's observation is significant because, in a counterintuitive fashion, it does not place the emphasis on the answers that are modified through a paradigm shift but on the transformation of traditional questions and problems into new and different ones. Like a shift in the light with which the puppets in a Chinese shadow play are illuminated, a paradigm shift allows us to see the figures in completely new shapes, forms and

movement. In this spirit, this dissertation tried to present a neuropolitical perspective on politics that aims to help political thinkers and cognitive scientists to view their figures of analysis anew, not just by giving novel answers about them, but most importantly, by transforming our imagination of which questions can be asked of the political brain.

BIBLIOGRAPHY

Abel, L. (1963). The aesthetics of evil: Hannah Arendt on Eichmann and the Jews. *Partisan Review, 30, 2*, 211-30.

Abelson, R. P., Kinder, D. R., Peters, M. D., & Fiske, S. T. (1982). Affective and semantic components in political person perception. *Journal of Personality and Social Psychology, 42*, 619–630.

Abrams, D., & Hogg, M. A. (1988). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. *European Journal of Social Psychology, 18*, 317-334.

Achen, C.H., & Bartels, L.M. (2016). Democracy for realists: Why elections do not produce responsive government. Princeton, NJ: Princeton University Press.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Adolphs, R. (2006). How do we know the minds of others? Domain-specificity, simulation, and enactive social cognition. *Brain Research, 1079*, 25-35.

Adorno, T. W. (2005). *Minima moralia: Reflections on a damaged life*. London: Verso.

Adorno, T. W. (1974). Theses against occultism. *Telos, 1974*, 7-12.

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., Nevitt-Sanford, R. (1950). *The authoritarian personality.* New York: Harper.

Akkerman, T., deLange, S.L., & Rooduijn, M. (2016). *Radical right-wing populist parties in Western Europe: Into the mainstream?*. New York: Routledge.

Albanis, E. (2002). *German-Jewish cultural identity from 1900 to the aftermath of the First World War : a comparative study of Moritz Goldstein, Julius Bab and Ernst Lissauer*. Tübingen : Niemeyer.

Alexievich, S. (2016). *Secondhand time: The last of the Soviets*. New York: Random House.

Alford, John R., Hibbing, John R. (2008). The New Empirical Biopolitics. *Annual Review of Political Science, 11*, 183-203.

Ali, T. (2002). *The clash of fundamentalists: Crusades, Jihads and modernity*. London: Verso.

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, Mass.: Addison-Wesley.

Almond, G.A., & Verba, S. (1963) *The civic culture: political attitudes and democracy in five nations*. Princeton, N.J.: Princeton University Press.

Altemeyer, B. (1981). *Right-wing authoritarianism*. Winnipeg: University of Manitoba Press.

Altemeyer, B. (2004 [1998]). The other "authoritarian personality". In J.T. Jost & J. Sidanius (Eds.), *Political psychology: key readings*. New York, Hove: Psychology Press.

Ames, Daniel L., Fiske, Susan T. (2010). Cultural neuroscience. *Asian Journal of Social Psychology, 13*, 2, 72-82.

Ames, D. L., & Fiske, S. T. (2013). Encountering the unexpected under outcome dependency: Power relations alter the neural substrates of impression formation. *NeuroImage, 83*, 599–608.

Amnesty International. (2005). *Human rights for human dignity: a primer on economic, social and cultural rights*. London, UK: Amnesty International Publications.

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science, 18*, 524–530.

Amodio, David M., Frith, Chris D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci, 7*, 268-277.

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *Journal of Personality and Social Psychology, 84*, 738-753.

Amodio, D. M, Jost, J. T., Master, Yee, S.L., & Cindy, M. (2007). Neurocognitive correlates of liberalism and conservatism. *Nature Neuroscience, 10*, 1246-1247.

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience, 1*, 26-36.

Andrighetto, L., Baldissari, C., Lattanzio, S., Loughnan, S., & Volpato, C. (2014). Human-itarian aid? Two forms of dehumanization and willingness to help after natural disasters. *British Journal of Social Psychology, 53*, 3, 573-584.

Appiah, K. A. (1997). The multiculturalist misunderstanding. *New York Review of Books 44*, 15, 36.

Aquinas, T. (2008) *On love and charity: readings from the Commentary on the sentences of Peter Lombard*. Washington, D.C.: Catholic University of America Press.

Arendt, H. (1951). *The origins of totalitarianism*. New York: Harcourt Brace.

Arendt, H. (1994[1963]). *Eichmann in Jerusalem: A report on the banality of evil*. New York: Penguin Books.

Arnhart, L. (1998). The new Darwinian naturalism in political theory. *Zygon, 33*, 369-393.

Ascher, W., & Mirovitskaya, N. (2015). *Development Strategies and Inter-Group Violence*: Insights on conflict-sensitive development. New York: Palgrave Macmillan.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*, 1311-1320.

Astington, J. W., & Baird, J. A. (Eds.). (2005). *Why Language Matters for Theory of Mind*. Oxford, UK: Oxford University Press.

Ausderan, J. (2014). How naming and shaming affects human rights perceptions in the shamed country. *Journal of Peace Research, 51*, 81-95.

Avis, J., & Harris, P. L. (1991). Belief-desire reasoning among Baka children: evidence for a universal conception of mind. *Child Development, 62*, 460-467.

Azize, J. (2007). Human dignity and euthanasia law. *University of Notre Dame Australia Law Review, 9*, 47-74.

Backman, C.R. (2012) *Cultures of the West: a combined history.* Oxford, UK: Oxford University Press.

Bacon, H.H. (1961). *Barbarians in Greek tragedy*. New Haven, Conn.: Yale University Press.

Bain, P.G., Vaes, J.V., & Leyens J.-P. (2014) *Humanness and Dehumanization*. New York and London: Routledge.

Bandura, A., Underwood, B., & Fromson, M.E (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality, 9*, 253-269.

Bantimaroudis, P. (2015). A mediated assessment of Samuel Huntington's 'Clash of Civilizations': The cultural framing hypothesis. *International Journal of Media and Cultural Politics, 11*, 1, 73-85.

Bar-Tal, Daniel (1989). *Stereotyping and Prejudice*. New York: Springer Nature.

Barceló, (2014). Re-examining a modern classic: Does Putnam's Making Democracy Work suffer from spuriousness? *Modern Italy, 19*, 4, 457-471.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*, 241-251.

Barry, B. (2001) *Culture and equality: an egalitarian critique of multiculturalism*. Cambridge, UK: Polity Press.

Bartlett, T. (2017, January 5) Can we really measure implicit bias? Maybe not. *Chronicle of Higher Education*. Retrieved from http://www.chronicle.com/article/Can-We-Really-Measure-Implicit/238807.

Bauer, J.R., & Bell, D.A. (Eds.) (1999). *The East Asian challenge for human rights*. New York: Cambridge University Press.

Baum, W. M. (2002). From molecular to molar: a paradigm shift in behavior analysis. *Journal of the Experimental Analysis of Behavior, 78*, 1, 95-116.

Bauman, C. W., McGraw, A., Bartels, P., Daniel M., & Warren, C. (2014). Revisiting external validity: concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass, 8*, 536-554.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*, 497-529.

Beck, U. (2006). *The cosmopolitan vision*. Cambridge; Malden, MA: Polity Press.

Beitz, C. (2009). *The idea of human rights*. Oxford, UK: Oxford University Press.

Bendor, J. (2010). *Bounded rationality and politics*. Berkeley, CA: University of California Press.

Bennett, M. R., & Hacker, P.M.S. (2003). *Philosophical foundations of neuroscience.* Hoboken, NJ: Wiley-Blackwell.

Bennett, M. R., & Hacker, P.M.S. (2012). *History of cognitive neuroscience*. Hoboken, NJ: Wiley-Blackwell.

Bentham, J. (1843). *Anarchical fallacies*. Retrieved from http://oll.libertyfund.org/titles/bentham-the-works-of-jeremy-bentham-vol-2.

Benton, T. (1991). Biology and social science: why the return of the repressed should be given a (cautious) welcome. *Sociology 25*, 1-30.

Benz, W. (1999). The Holocaust: a German historian examines the genocide. New York: Columbia University Press.

Berg, B.J. (2009). *Sexism in America: Alive, well and ruining our future*. Chicago, IL: Lawrence Hill.

Bergsieker, H. B., Shelton, J. N., & Richeson, J. A. (2010). To be liked versus respected: Divergent goals in interracial interactions. *Journal of Personality and Social Psychology, 99,* 248–264.

Berkley, G. (1998[1710]). *A treatise concerning the principles of human knowledge*. Oxford, UK: Oxford University Press.

Berlin, I. (2013). *Against the current: Essays in the history of ideas*. Princeton, NJ: Princeton University Press.

Berman, S. (2003). Islamism, revolution, and civil society. *Perspectives on Politics, 1*, 2, 257-272.

Bernasconi, R., & Mann, A.M. (2005). The contradictions of racism: Locke, slavery, and the two treaties. In A. Valls (Ed.), *Race and racism in modern philosophy*. Ithaca, NY: Cornell University Press.

Berntson, G. G., Sarter, M., & Cacioppo, J. T. (1998). Anxiety and cardiovascular reactivity: the basal forebrain cholinergic link. *Behavioural Brain Research, 94*, 225-248.

Beyer, C. (2015) Edmund Husserl. *Stanford Encyclopedia of Philosophy*. Retrieved at https://plato.stanford.edu/entries/husserl/.

Bialik, C. (2015, May 20). As a major retraction shows, we're all vulnerable to faked data. *FiveThirtyEight Blog*. Retrieved from https://fivethirtyeight.com/datalab/as-a-major-retraction-shows-were-all-vulnerable-to-faked-data/.

Blakemore, S.-J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: where are we heading? *Trends in Cognitive Sciences, 8*, 216-222.

Blanco, E.I. (2014, December 11). Intolerance, human dignity and rights. *Campaign Action Blog*. Retrieved from http://blog.nohatespeechmovement.org/intolerance-human-dignity-and-human-rights/.

Blank, R., & Hines, S.M. (2001). *Biology and political science*. London, New York: Routledge.

Blass, T. (1999). The Milgram paradigm after 35 years: some things we now know about obedience to authority. *Journal of Applied Social Psychology, 29*, 955-978.

Bohannon, J. (2015, May 28) *Science* retracts gay marriage paper without agreement of lead author LaCour. *Science Online*. Retrieved from http://www.sciencemag.org/news/2015/05/science-retracts-gay-marriage-paper-without-agreement-lead-author-lacour.

Borders, A., Earleywine, M., & Jajodia, A. (2010). Could mindfulness decrease anger, hostility, and aggression by decreasing rumination? *Aggressive Behavior, 36*, 1, 28-44. doi:10.1002/ab.20327

Boski, P. (2008). Five meanings of integration in acculturation research. *International Journal of Intercultural Relations, 32*, 142-153.

Brewer, M. B. (1999). The Psychology of Prejudice: Ingroup Love and Outgroup Hate? *Journal of Social Issues, 55*, 429-444.

Brewer, M.B., & Miller, N. (1996). *Intergroup relations*. Buckingham: Open University Press.

Broockman, D., Kalla, J., & Aronow, P. (2015, May 19). *Irregularities in LaCour (2014)*. Online paper. Retrieved from https://stanford.edu/~dbroock/broockman_kalla_aronow_lg_irregularities.pdf.

Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience, 1*, 27-51.

Brown, G.W., & Held, D. (2010). *The cosmopolitanism reader*. Cambridge, UK: Polity Press.

Browning, C. (1992). Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland. New York: Harper Perennial.

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Brustein, W. (2003). Roots of hate: anti-semitism in Europe before the Holocaust. Cambridge, UK: Cambridge University Press.

Buchanan, A.E. (2013). *The heart of human rights*. Oxford, UK: Oxford University Press.

Buchanan, A.E. (2015). Why International legal human rights? In R.Cruft, S.M. Liao, & M. Renzo (Eds.), *Philosophical foundations of human rights* (pp.244-263). Oxford, UK: Oxford University Press.

Bush, G.H.W. (1992, August 6). Remarks on the situation in Bosnia and an exchange with reporters in Colorado Springs. *Public Papers of the President of the United States, 2*, 1315-1318.

Bush, G.W. (2001, September 15). Remarks in a meeting with the National Security team and an exchange with reporters at Camp David, Maryland. *Public Papers of the President of the United States, 2*, 1111-1113.

Cacioppo, J. T. (1994). Social neuroscience: Automatic, neuroendocrine, and immune responses to stress. *Psychophysiology, 31*, 113-28.

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist, 47*, 1019-1028.

Cacioppo, J. T., Crites, S. L., Jr., & Gardner, W. L. (1996). Attitudes to the right: Evaluative processing is associated with lateralized late positive event related brain potentials. *Personality and Social Psychology Bulletin, 22*, 1205 – 1219.

Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology, 85*, 650 – 661.

Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2008). Social neuroscience: bridging social and biological systems. In C. Sansone, C.C. Morf, & A.T. Panter (eds.), *The SAGE Handbook of Methods in Social Psychology* (pp. 383-404). Thousand Oaks: Sage Publications.

Cacioppo, J.T., Visser, P.S., & Pickett, C.L. (2006). *Social neuroscience: people thinking about people.* Cambridge, Mass.: MIT Press.

Cacioppo, J. T., Cacioppo, S., & Gollan, J. K. (2014). The negativity bias: conceptualization, quantification, and individual differences. *Behavioral Brain Sciences, 37*, 309-310.

Cacioppo, J. T., & Visser, P. S. (2003). Political Psychology and Social Neuroscience: Strange Bedfellows or Comrades in Arms?. *Political Psychology, 24*, 647-656.

Calvin, J. (1960). *Institutes of the Christian religion*, *vol. I.* Philadelphia: Westminster Press.

Cameron, C. D., Harris, L. T., & Payne, B. K. (2016). The emotional cost of humanity: anticipated exhaustion motivates dehumanization of stigmatized targets. *Social Psychological and Personality Science, 7*, 105-112.

Capestany, B. H., & Harris, L. T. (2014). Disgust and biological descriptions bias logical reasoning during legal decision-making. *Social Neuroscience, 9*, 265-277.

Carter, T.J., Ferguson, M.J., & Hassin, R.R. (2011). Implicit nationalism as systems justification: the case of the United States of America. *Social Cognition, 29*, 3, 341-359.

Castano, E., & Giner-Sorolla, R. (2006). Not quite human: infrahumanization in response to collective responsibility for intergroup killing.. *Journal of Personality and Social Psychology, 90*, 804-818.

Castelli, L., & Carraro, L. (2011). Ideology is related to basic cognitive processes involved in attitude formation. *Journal of Experimental Social Psychology, 47*, 1013-1016.

Castells, M. (1997). *The power of identity (The information age: Economy, society, and culture, vol.II)*. Malden, MA: Blackwell.

Cavalieri, P. (2001). *The animal question: why nonhuman animals deserve human rights*. Oxford, UK: Oxford University Press.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chalmers, D. (1996). *The conscious mind.* Oxford, UK: Oxford University Press.

Chan, C.K. (2005). *Human dignity and welfare systems*. Bristol, UK: Policy Press.

Charney, E. (2012). Candidate genes and political behavior. *American Political Science Review 106*, 1, 1-34.

Charney, E. (2014). Conservatives, liberals, and "the negative". *Behavioral Brain Sciences, 37*, 3, 310-311.

Chartrand, T., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: nonconscious goal priming reproduces effects of explicit task instructions. *Journal of Personality and Social Psychology, 71*, 464-478.

Chernoff, F. (2009). The ontological fallacy: a rejoinder on the status of scientific realism in International Relations. *Review of International Studies, 35*, 2, 371-395.

Chiao, J.Y., & Ambady, N. (2007). Cultural neuroscience: Parsing universality and diversity across levels of analysis. In S. Kitayama & D. Cohen (Eds.), *Handbook of Cultural Psychology*. New York: Guilford Press.

Chiao, J.Y., Harada, T., Komeda, H., & et al. (2009). *Dynamic cultural influences on neural representations of self.* Journal of Cognitive Neuroscience, 22, 1, 1-11.

Chirot, D. (2001). A clash of civilizations or of paradigms?. *International Sociology, 16*, 341-360.

Chomsky, N., & Rajchman, J. (2006). *The Chomsky-Foucault debate: on human nature*. New York: New Press.

Choudhury, S., & Slaby, J. (Eds.). (2011). *Critical Neuroscience*. Chichester, West Sussex: Wiley-Blackwell.

Christie, R., & Geis, F. (1970). *Studies in Machiavellianism*. New York: Academic Press.

Churchland, P. S. (1981). On the alleged backwards referral of experiences and its

relevance to the mind-body problem. *Philosophy of Science, 48*, 165-181.

Cicero, M.T. (2014). *De officiis*. Cambridge, MA: Harvard University Press.

Cikara, M., Botvinick, M. M., & Fiske, S. T. (2011). Us versus them. *Psychological Science, 22*, 306-313.

Cikara, M., E. Bruneau, J.J. Van Bavel, & R. Saxe (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology, 55*, 110-125.

Claassen, C. (2016). Group entitlement, anger and participation in intergroup violence. *British Journal of Political Science, 46*, 1, 127-148.

Clark, R. (1970). Racism and human dignity. *Journal of Religion and Health, 9*(4), 316-322.

Clark, G. (2016). Augustine and the merciful barbarians. In D. Shanzer & R.W. Mathisen (Eds.), *Romans, barbarians, and the transformation of the Roman World* (pp.33-42). Farnham; Burlington: Ashgate.

Cloutier, J., Gabrieli, J. D., O'Young, D., & Ambady, N. (2011). An fMRI study of violations of social expectation: when people are not who we expect them to be. *Neuroimage, 57*, 583–588.

Cohen, G.I. (2015). This is your brain on human rights: moral enhancement and human rights. *The Law & Ethics of Human Rights, 9*, 1, 1-41.

Coleman, M. G. (2005). Racism in academia: The white superiority supposition in the "unbiased" search for knowledge. *European Journal of Political Economy, 21*, 762-774.

Colman, A.M. (2006). *A dictionary of psychology.* Oxford, UK: Oxford University Press.

Commonwealth Human Rights Initiative (2011). *Annual report – headquarters India*. Retrieved from http://www.humanrightsinitiative.org/publications/areport/2012/CHRIAnnualReport2011-12.pdf.

Connolly, W. (2002). *Neuropolitics: thinking, culture, speed.* Minneapolis: University of Minnesota Press.

Cooper, J. B., & Siegel, H. E. (1956). The Galvanic Skin Response as a Measure of Emotion in Prejudice. *The Journal of Psychology, 42*, 149-155.

Cory, G. A. (2004). *The consilient brain: the bioneurological basis of economics, society and politics*. New York: Kluwer Academic.

Crawford, N. C. (2009). Human nature and world politics: rethinking 'man'. *International Relations, 23*, 271-288.

Cruft, R., Liao, S. M., & Renzo, M. (Eds.). (2015). *Philosophical foundations of human rights*. Oxford, UK: Oxford University Press.

Cuddy, A.J.C., Fiske, S.T., Kwan, V.S.Y., Glick, P., Demoulin, S., & et al. (2009). Stereotype content model across cultures: towards universal similarities and

differences. *British Journal of Social Psychology, 48*, 1-33.

Cunningham, W. A., & Van Bavel, J. J. (2009). A neural analysis of intergroup perception and evaluation. In G.G. Berntson & J. T. Cacioppo (Eds.), *Handbook of neuroscience for the behavioral sciences* (Vol. 1, pp. 975–984). Hoboken, NJ: Wiley.

D'Esposito, M., Ballard, D., Aguirre, G., & Zarahn, E. (1998). Human Prefrontal Cortex Is Not Specific for Working Memory: A Functional MRI Study. *NeuroImage, 8*, 274-282.

Dalsklev, M., & Kunst, J.R. (2015) The effect of disgust-eliciting media portrayals on outgroup dehumanization and support of deportation in a Norwegian sample, *International Journal of Intercultural Relations, 47*, 28-40.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science, 264*(5162), 1102-1105.

Dancygier, R.M. (2010). *Immigration and conflict in Europe*. New York: Cambridge University Press.

Dardenne, B., Dumont, M., Sarlet, M., Phillips, C., Balteau, E., & et al. (2013). Benevolent sexism alters executive brain responses. *Neuroreport, 24,* 10, 572-577.

Dawkins, R. (1976). *The selfish gene.* New York: Oxford University Press.

Decety, J., & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience, 2*, 1-24.

Degler, C. (1991). *In search of human nature: The decline and revival of Darwinism in American social thought*. New York: Oxford University Press.

Demoulin, S., Vaes, J., Gaunt, R., Paladino, M.P., & Leyens, J.-P. (2007). Infra-humanization: The wall of group difference. *Social Issues and Policy Review, 1*, 1, 139-172.

Dennett, D.C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Derks, B., Scheepers, D., & Ellemers, N. (Eds.), (2013). *Neuroscience of prejudice and intergroup relations*. London, UK: Psychology Press.

Descartes, R. (1996[1641]). *Meditations on first philosophy*. Cambridge, UK: Cambridge University Press.

Deutsch, M., & Kinnvall, C. (2002) What is political psychology?. In K.R. Monroe (Ed.), *Political Psychology*. Mahwah, NJ: Erlbaum.

DeWall, C. N., & Baumeister, R. F. (2006). Alone but feeling no pain: effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *Journal of Personality and Social Psychology, 91*, 1-15.

DiCosmo, N. (2002). *Ancient China and its enemies: The rise of nomadic power in East Asian history*. Cambridge, UK: Cambridge University Press.

Dobbs, M. (1995, December 3). Bosnia crystallizes U.S. post-Cold War role. *The Washington Post*, A1.

Dodd, M. D., Balzer, A., Jacobs, C.M., Gruszczynski, M.W., Smith, & et al. (2012). The political Left rolls with the good and the political Right rolls with the bad: connecting physiology and cognition to preferences. *Philosophical Transactions of the Royal Society of London 367*, 1589, 640–49.

Donnelly, J. (2003). *Universal human rights in theory and practice*. Ithaca and London: Cornell University Press.

Donnelly, J. (1998). Human Rights: A New Standard of Civilization?. *International Affairs, 74*, 1-23.

Dovidio, J. F., Eller, A., & Hewstone, M. (2011). Improving intergroup relations through direct, extended and other forms of indirect contact. *Group Processes & Intergroup Relations, 14*, 147-160.

Dovidio, J. F., Gaertner, S. L., & Validzic, A. (1998). Intergroup bias: Status, differentiation, and a common in-group identity. *Journal of Personality and Social Psychology, 75*, 109-120.

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive Racism and Selection Decisions: 1989 and 1999. *Psychological Science, 11*, 315-319.

Dovidio, J. F., Gaertner, S. E., Kawakami, K., & Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural Diversity and Ethnic Minority Psychology 8*, 2, 88-102.

Dovidio, J.F., & Gaertner, S.L. (2005). Understanding and addressing contemporary racism: From aversive racism to the common ingroup identity model. *Journal of Social Issues, 61*, 3, 615-639.

Dretske, F. (1969). *Explaining behavior: reasons in a world of causes.* Cambridge, Mass.: The MIT Press.

Druckman, J.N., Green, D.P., Kuklinski, J.H., & Lupia, A. (2006). The growth and development of experimental research political science. *American Political Science Review 100*, 627-36.

DuBois, W.E.B. (1897). *The conservation of races*. Washington, DC: The American Negro Academy.

Durante, F., S. T. Fiske, N. Kervyn, and A.J.C. Cuddy, et al. (2013). Nations' income inequality predicts ambivalence in stereotype content: how societies mind the gap. *British Journal of Social Psychology, 52*, 726-746.

Easton, D. (1965). *A framework for political analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Eisenberg, A., & Spinner-Halev, J. (Eds.) (2005). *Minorities within minorities: equality, rights, and diversity.* Cambridge, UK: Cambridge University Press.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 7900-7905.

Ellison, R. (1997) *Invisible man.* New York: Chelsea House Publishers.

Elster, J. (1993). *Political psychology*. Cambridge, UK: Cambridge University Press.

European Commission (2013). *Divercities: creating social cohesion, social mobility and economic performance in today's hyper-diversified cities*. Project Brief. Retrieved from https://www.urbandivercities.eu/about-divercities.

Eysenck, H. J., & Eysenck, S.B. (1976). *Psychoticism as a dimension of personality.* New York: Crane, Russak, & Company.

Ezorsky, G. (1966). Hannah Arendt answered. *Dissent Vol. XIII, 2*, 172-82.

Faingold, C. L., & Blumenfeld, H. (2014). *Neuronal Networks in Brain Function, CNS Disorders, and Therapeutics*. Amsterdam: Elsevier.

Fanon, F. (1967). *Black skin, white masks*. New York: Grove Press.

Farer, T. J. (2000). Restraining the barbarians: Can International criminal law help?. *Human Rights Quarterly, 22*, 90-117.

Faro, S. H., & Mohamed, F. B. (Eds.). (2006). *Functional MRI*. New York: Springer Nature.

Fausto-Sterling, A. (1992). *Myths of Gender: Biological Theories about Women and Men.* New York: Basic Books.

Fine, C. (2010). *Delusions of Gender: How Our Minds, Society and Neurosexism Create Difference.* London: Icon Books.

Finkelstein, G. (2014). Emil du Bois-Reymond on "the seat of the soul". *Journal of the History of the Neurosciences, 23*, 1, 45-55.

Fiske, S., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences, 111*, 4, 13593-13597.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences, 11*, 77-83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878-902.

Fiske, S.T., Lin, M., & Neuberg, S.L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231-254). New York: Guilford Press.

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & et al. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57*, 109-128.

Flikschuh, K., & Ypi, L. (Eds.) (2014). *Kant and colonialism: historical and critical perspectives*. Oxford, UK: Oxford University Press.

Fodor, J.A. (1975). *Psychosemantics*. Cambridge, Mass: The MIT Press.

Fowler, J. H., & Schreiber, D. (2008). Biology, politics, and the emerging science of human nature. *Science, 322*, 912-914.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 9673-9678.

Franklin, R.G., Stevenson, M.T., Ambady, N., & Adams Jr., R.B. (2015). Cross-cultural reading in the mind in the eyes and its consequences for International relations. In J.E. Warnick & D. Landis (Eds.), *Neuroscience in intercultural contexts* (pp.117-141). New York: Springer New York

Franklin, J. C. (2008). Shame on you: The impact of human rights criticism on political repression in Latin America. *International Studies Quarterly, 52*, 187-211.

Franks, D.D., & Smith, T.S (Eds.) (1999). *Mind, brain, and society: toward a neurosociology of emotion*. Stanford, CT: JAI Press.

Fraser, N. (2000). Rethinking recognition. *New Left Review, 3*, 3, 107-120

Fredrickson, B.L., Cohn, M.A., Coffey, K.A., Pek, J., and Finkel, S.M. (2008). Open hearts build lives: positive emotions, induced through loving kindness meditation, build consequential personal resources. *Journal of Personality and Social Psychology, 95*, 1045–1062.

Freeman, S.R. (2007). *Rawls*. London, New York: Routledge.

Frith, U. (2001). Mindblindness and the brain in autism. *Neuron, 32*, 969-979.

Frith, U, & Frith, C.D. (1999). Cognitive psychology – interacting minds – a biological basis. *Science, 286*, 5445, 1692-1695.

Frith, U., & Frith, C.D. (2005). Theory of mind. *Current Biology, 15*, 17, p.R644.

Fukuyama, F. (1992). *The end of history and the last man.* New York: Free Press.

Fukuyama, F. (2011). *The origins of political order: From prehuman times to the French revolution*. New York: Farrar, Strauss and Giroux.

Fukuyama, F. (2014). *Political order and political decay: From the Industrial revolution to the globalization of democracy*. London: Profile Books.

Galeotti, A. (2002). *Toleration as recognition*. Cambridge, UK: Cambridge University Press.

Gans, H.J. (1995). *The war against the poor: The underclass and antipoverty policy*. New York, NY: Basic Books.

Gaus, G. (1964, October 28). Günter Gaus im Gespräch mit Hannah Arendt. Television Interview Script (German original). Retrieved from http://www.rbb-online.de/zurperson/interview_archiv/arendt_hannah.html.

Gauthier, D. (1974). Rational Cooperation. *Noûs, 8*, 53-65.

Gianaros, P.J., Marsland, A.L., Sheu, L.K., Erickson, K.I., & Verstynen, T.D. (2013). Inflammatory pathways link socioeconomic inequalities to white matter architecture. *Cerebral Cortex, 23*, 9, 2058-2071.

Gibbons, F., O'Hara, R.E., Stock, M.L., Gerrard, M., Weng, C.-Y., & Wills, T. (2012).

The erosive effects of racism: reduced self-control mediates the relation between perceived racial discrimination and substance use in African American adolescents. *Journal of Personality and Social Psychology, 102*, 5, 1089-1104.

Glendon, M.A. (2001). *A world made new: Eleanor Roosevelt and the Universal Declaration of Human Rights*. New York: Random House.

Glimcher, P.W., & Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science, 306*, 5695, 447-452.

Goff, P.A., Eberhardt, J.L., Williams, M.J., & Jackson, M.C. (2008). Not yet human: implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology, 94*, 2, 292-306.

Goffart, W. (2006). *Barbarian tides: The migration and the later Roman Empire*. Philadelphia, PA: University of Pennsylvania Press.

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience, 4*, 8, 845-850.

Gowans, C.W. (2001). Kant's impure ethics. *International Philosophical Quarterly, 41*, 3, 363-369.

Gozzi, M., Zamboni, G., Krueger, F., & Grafman, J. (2010). Interest in politics modulates neural activity in the amygdala and ventral striatum. *Human Brain Mapping, 31*, 1763–1771.

Graham, G. (1984). Sensation and classification. *Behavioral and Brain Sciences, 7*, 558-559.

Graham, G. (2015). Behaviorism. *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/behaviorism.

Green, D.P., & Shapiro, I. (1994). *Pathologies of rational choice theory: a critique of applications in political science*. New Haven, Conn.: Yale University Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464 – 1480.

Gross, C.G., & Rocha-Miranda, C.E., & Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35*, 96-111.

Gunnell, J. G. (2013). The reconstitution of political theory: David Easton, behavioralism, and the long road to system. *Journal of the History of Behavioral Sciences, 49*, 190-210.

Gurr, T. R. (1970). *Why men rebel.* Princeton, NJ: Princeton University Press.

Gutchess, A. H., & Goh, J. O. S. (2013). Refining concepts and uncovering biological mechanisms for cultural neuroscience. *Psychological Inquiry, 24*, 31-36.

Habermas, J. (1981). *Theory of communicative action*. Boston: Beacon Press.

Habermas, J. (2010). The concept of human dignity and the realistic utopia of human rights. *Metaphilosophy, 41*, 4, 464-480.

Hafner-Burton, E. M. (2008). Sticks and stones: Naming and shaming the Human rights enforcement problem. *International Organization, 62*, 689-716.

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-Group Love" and "Out-Group Hate" as motives for individual participation in intergroup conflict: a new game paradigm. *Psychological Science, 19*, 405-411.

Hall, M., & Jackson, P.T. (2007). Introduction: Civilizations and International relations theory. In M. Hall & P.T. Jackson (Eds.), *Civilizational identity: The production and reproduction of "civilizations" in International relations* (pp.1-14). Basingstoke; New York: Palgrave Macmillan.

Han, S., Mao, L., Gu, X., Zhu, Y., Ge, Y. & et al. (2008). Neural consequences of religious belief on self-referential processing. *Social Neuroscience, 3*, 1, 1-15.

Han, S. (2015). Understanding cultural differences in human behavior: a cultural neuroscience approach. *Current Opinion in Behavioral Sciences, 3*, 68-72.

Haney, C., Banks, C., & Zimbardo, P. (1973). A study of prisoners and guards in a simulated prison. Washington, DC: Naval Research Academy.

Hansen, R., & King, D. (2013). *Sterilized by the state: eugenics, race, and the population scare in Twentieth-Century North America.* Cambridge, UK: Cambridge University Press.

Hansen, L. (2006). *Security as practice: Discourse analysis and the Bosnian war.* London; New York: Routledge.

Haque, O.S., & Waytz, D.M. (2012). Dehumanization in medicine: causes, solutions, and functions. *Perspectives on Psychological Science, 7*, 2, 176-186.

Harlow, J. M. (1868). Recovery from the passage of an iron rod through the head. *Publications of the Massachusetts Medical Society*, *2*, 327–347.

Harmon-Jones, E., & Winkielman, P. (2007). *Social neuroscience: integrating biological and psychological explanations of behavior*. New York: The Guildford Press.

Harrell, J. P., Hall, S., & Taliaferro, J. (2003). Physiological responses to racism and discrimination: an assessment of the evidence. *American Journal of Public Health, 93*, 243-248.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low. *Psychological Science, 17*, 847-853.

Harris, L. (2017). *Invisible mind: Flexible social cognition and dehumanization*. Cambridge, Mass: MIT Press.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low. *Psychological Science, 17*, 847-853.

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in the mPFC. *Social and Cognitive Affective Neuroscience, 2*, 45-51.

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised

perception. *European Review of Social Psychology, 20*, 192-231.

Harris, L. T., & Fiske, S. T. (2011). Dehumanized perception: a psychological means to facilitate atrocities, torture, and genocide? *Zeitschrift für Psychologie, 219*, 175-181.

Harris, L. T., & Fiske, S. T. (2008). The Brooms in Fantasia: Neural Correlates of Anthropomorphizing Objects. *Social Cognition, 26*, 210-223.

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *NeuroReport, 11*, 2351–2355.

Hartleb, F. (2011). After their establishment: right-wing populist parties in Europe. *European View, 10*, 2, 267-268.

Haslam, N. (2006). Dehumanization: An Integrative Review. *Personality and Social Psychology Review, 10*, 252-264.

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology, 65*, 399-423.

Hatemi, P. K., & McDermott, R. (2011). The Normative Implications of Biological Research. *PS: Political Science and Politics, 44*, 325-329.

Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience, 3*, 142-151.

Heflick, N. A., & Goldenberg, J. L. (2009). Objectifying Sarah Palin: Evidence that objectification causes women to be perceived as less competent and less fully human. *Journal of Experimental Social Psychology, 45,* 598–601.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales: The reference-group problem. *Journal of Personality and Social Psychology, 82*, 903–918.

Heppner, W., Cascio, E.V., Campbell, W.K., Goldman, B.M., Lakey, C.E., & et al. (2008). Mindfulness as a means of reducing aggressive behavior: dispositional and situational evidence. *Aggressive behavior, 34*, 5, 486-496.

Herbert, R.T. (1998). Dualism/materialism. *Philosophical Quarterly, 48*, 159-175.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology, 53*, 575-604.

Hibbing, J. R., Smith, K. B., Alford, J. R. (2014). Differences in negativity bias underlie variations in political ideology. *Behav Brain Sci, 37*, 297-307.

Hibbing, J. R., & Smith, K. B. (2007). The Biology of Political Behavior: An Introduction. *The ANNALS of the American Academy of Political and Social Science, 614*, 6-14.

Hibbing, J. R., Smith, K. B., & Alford, J. R. (2014). Differences in negativity bias underlie variations in political ideology. *Behavioral Brain Sciences, 37*, 297-307.

Hill, J.H. (2008). *The everyday language of white racism*. Hoboken, NJ: Wiley-Blackwell.

Hill, J.D. (2009). *Beyond blood identites: post-humanity in the 21st century*. Lanham, MD: Lexington Books

Hillygus, D.S., & Shields, T.G. (2008). *The persuadable voter: Wedge issues in presidential campaigns*. Princeton, NJ: Princeton University Press.

Hines, S.M., Funke, O., Losco, J., Stewart, P., Blank, R. H. (2014). *Politics and the life sciences: the state of the discipline.* Bingley: Emerald Group Publishing Limited.

Hobbes, T. (2005[1651]). *Leviathan*. Cambridge, UK: Cambridge University Press.

Hobson, J.M. (2004). *The Eastern origins of Western civilization*. Cambridge, UK: Cambridge University Press.

Hodgson, D. (1988). *The Mind Matters: Consciousness and Choice in a Quantum World*. Oxford, UK: Oxford University Press.

Hodson, G., & Costello, K. (2007). Interpersonal Disgust, Ideological Orientations, and Dehumanization as Predictors of Intergroup Attitudes. *Psychological Science, 18*, 691-698.

Hodson, G., MacInnis, C., & Costello, K. (2014). (Over)valuing "humanness" as an aggravator of intergroup prejudice and discrimination. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp.86-110). New York and London: Routledge.

Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences of the United states of America, 113*, 4296-4301.

Honneth, A. (1996). *The struggle for recognition: the moral grammar of social conflicts*. Cambridge, Mass.: The MIT Press.

Hornsey, M. J., Harth, N.S., & Barlow, F.K. (2008). Emotional responses to rejection of gestures of intergroup reconciliation. *Personality and Social Psychology Bulletin, 37*, 6, 815-829.

Huddy, L., & Feldman, S. (2014). Not so simple: The multidimensional nature and diverse origins of political ideology. *Behavioral Brain Sciences, 37*, 312-313.

Huddy, L., Feldman, S., & Lown, P. (2014). When empathy succeeds and fails: public support for social welfare policies. *Online Working Paper*. Retrieved from http://lore.gu.se/digitalAssets/1525/1525704_feldman-huddy_empathy.pdf

Huddy, L., Sears, D. O., & Levy, J. S. (Eds.). (2013). *The Oxford Handbook of Political Psychology*. Oxford, UK: Oxford University Press.

Hughes, J., & Churchland, P. S. (1995). My behavior made me do it: The uncaused cause of teleological behaviorism. *Behavioral Brain Sciences, 18*, 130.

Hughes, C.R. (2006). *Chinese nationalism in the global era*. London; New York: Routledge.

Human Rights Watch (2009). *Broken system: dysfunction, abuse, and impunity in the Indian Police.* Annual Report. Retrieved from

https://www.hrw.org/report/2009/08/04/broken-system/dysfunction-abuse-and-impunity-indian-police.

Hume, D. (1975[1739]). *The Clarendon Edition of the Works of David Hume: A Treatise of Human Nature, Vol. 1: Texts*. Oxford, UK: Oxford University Press.

Humphreys, M., & Jacobs, A. M. (2015). Mixing methods: A Bayesian approach. *The American Political Science Review, 109*, 4, 653-673.

Humphreys, M., & Weinstein, J. M. (2009). Field experiments and the political economy of development. *Annual Review of Political Science, 12*, 367-378.

Huntington, S.P. (1996). *The clash of civilizations and the remaking of world order.* New York: Simon & Schuster.

Huntington, S. (2004). *Who are we?: The challenges to America's national identity*. New York: Simon & Schuster.

Husserl, E. (1999[1931]). *Cartesian meditations: an introduction to phenomenology*. Dordrecht: Springer Netherlands.

Husserl, E. (1950). *Husserliana: Gesammelte Werke* [Husserliana: Collected works]. Den Haag: M. Nijhoff.

Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J. C., & et al. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526-2528.

Ihrig, S. (2016). *Justifying genocide: Germany and the Armenians from Bismarck to Hitler*. Cambridge, Mass.: Harvard University Press.

Insel, T. R., & Winslow, J. T. (1998). Serotonin and neuropeptides in affiliative behaviors. *Biological Psychiatry, 44*, 207-219.

Jackson, P.T. (2011). *The conduct of inquiry in International relations: philosophy of science and its implications for the study of world politics*. London and New York: Routledge.

Jackson, P.T. (2003). Defending the West: Occidentalism and the formation of NATO. *The Journal of Political Philosophy, 11*, 3, 223-252.

Jamal, A. (2012). *Of empires and citizens: Pro-American democracy or no democracy at all?* Princeton, NJ: Princeton University Press.

Janoff-Bulman, R., & Carnes, N. C. (2014). Motivation and morality: Insights into political ideology. *Behavioral Brain Sciences, 37*, 316-317.

Jervis, R. (1976). *Perception and misperception in International politics.* Princeton, NJ: Princeton University Press.

Jervis, R. (1997). *Systems effects: complexity and the analysis of political and social life*. Princeton, NJ: Princeton University Press.

Jezzard, P., Matthews, P. M., & Smith, S. M. (Eds.). (2001). *Functional Magnetic Resonance Imaging*. Oxford, UK: Oxford University Press.

Johnson, D.M., & Erneling, C.E. (1997). *The future of the cognitive revolution.* New York: Oxford University Press.

Johnston, D. (1986). *The rhetoric of Leviathan: Thomas Hobbes and the politics of cultural transformation*. Princeton, NJ: Princeton University Press

Jones, B.D. (1999). Bounded rationality. *Annual Review of Political Science, 2*, 1, 297-321.

Jones, B.G. (2006). Introduction: International relations, eurocentrism, and imperialism. In B.G. Jones (Ed.), *Decolonizing International Relations* (pp.1-22). Lanham: Rowman & Littlefield.

Jost, J. T., & Sidanius, J. (Eds.), (2004*). Political psychology: Key readings*. New York, NY: Psychology Press/Taylor & Francis.

Jost, J. T., Glaser, J., Kruglanski, A.W., & Sulloway, F. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*, 339–375.

Jost, J. T., Nam, H. H., Amodio, D. M., & Van Bavel, J. J. (2014). Political neuroscience: the beginning of a beautiful friendship. *Advances in Political Psychology, 35*, 1, 3-42.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*, 263.

Kanai, R., Firth, C., Feilden, T., & Rees, G. (2011). Political orientations are correlated with brain structure in young adults. *Current Biology, 21*, 8, 677-680

Kaplan, J. T., Freedman, J., & Iacoboni, M. (2007). Us versus them: Political attitudes and party affiliation influence neural response to faces of presidential candidates. *Neuropsychologia, 45*, 55–64.

Kaplan, R.D. (1993). *Balkan ghosts: A journey through history*. New York: Vintage.

Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., Baron-Cohen, S. (1995). Is there a social module? language, face processing, and Theory of Mind in individuals with Williams syndrome. *Journal of Cognitive Neuroscience, 7*, 196-208.

Kashima, Y., & Margetts, E. A. (2014). On human-nature relationships. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp. 294-322). New York and London: Routledge.

Kateb, G. (2011). *Human dignity*. Cambridge, Mass.: Belknap Press.

Kaul, C., Ratner, K. G., & Van Bavel, J. J. (2013). Dynamic representations of race: Processing goals shape race decoding in the fusiform gyri. *Social Cognitive and Affective Neuroscience, 9*, 3, 326-332.

Keestra, M. (2012). Bounded mirroring: joint action and group membership in political theory and cognitive neuroscience. In F. Vander Valk (ed.), *Essays on neuroscience and political theory* (pp. 222-248). London and New York: Routledge.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology: A systematic text in the science of behavior.* New York: Appleton-Century-Crofts.

Kelman, H. G. (1973). Violence without Moral Restraint: Reflections on the Dehumanization of Victims and Victimizers. *Journal of Social Issues, 29*, 4, 25-61.

Kelman, H.C. (1989). *Crimes of obedience: toward a social psychology of authority and responsibility.* New Haven, Conn.: Yale University Press.

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues, 29*, 25-61.

Keyser, S.J., Miller, G.A., and Walker, E. (1978). Cognitive science in 1978. *An unpublished report submitted to the Alfred P. Sloan Foundation, New York*.

Kihlstrom, J. F. (2006). Does neuroscience constrain social-psychological theory? *Dialogue, 21*, 16-17.

Kitayama, S. (2002). Culture and basic psychological processes - toward a system view of culture: Comment on Oyserman et al. (2002). *Psychological Bulletin, 128*, 89-96.

Kleingeld, P. (2012). *Kant and cosmopolitanism: The philosophical ideal of world citizenship*. Cambridge, UK: Cambridge University Press.

Kliemann, D., Young, L., Scholz, J., & Saxe, R. (2008). The influence of prior record on moral judgment. *Neuropsychologia, 46*, 2949-2957.

Klotz, A., & Lynch, C. (2007). *Strategies for research in constructivist International Relations*. Armonk, NY: M.E. Sharpe.

Knutson, K. M., Wood, J. N., Spampinato, M. V., & Grafman, J. (2006). Politics on the brain: An fMRI investigation. *Social Neuroscience, 1*, 25-40.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature, 446*, 7138, 908-11.

Konner, M. (2002). *The tangled wing*. New York: Viking Compass.

Koppelman, A. (1996). *Antidiscrimination law and social equality*. New Haven, Conn.: Yale University Press.

Krain, M. (2012). *J'accuse!* Does naming and shaming perpetrators reduce the severity of genocides or politicides? *International Studies Quarterly, 56*, 3, 574-589.

Krendl, A. C., Macrae, C. N., Kelley, W. M., Fugelsang, J. A., & Heatherton, T. F. (2006). The good, the bad, and the ugly: An fMRI investigation of the functional anatomic correlates of stigma. *Social Neuroscience, 1*, 5-15.

Kretzmer, D., & Klein, E. (Eds.) (2002). *The concept of human dignity in human rights discourse*. The Hague; New York: Kluwer Law International.

Kteily, N., Bruneau, E., & Hodson, G. (2016). They see us as less than human: Metadehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology, 110*, 3, 343-370.

Kteily, N. , Bruneau, E., Waytz, A., & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology, 109*, 5, 901-931.

Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature Neuroscience, 15*, 940–948.

Kuhn, T. (1970[1962]). *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Kymlicka, W. (1995). *Multicultural citizenship: A liberal theory of minority rights*. Oxford; New York: Clarendon Press.

Kymlicka, W. (1998). *Finding our way: Rethinking ethnocultural relations in Canada*. Toronto: Oxford University Press.

Lacour, M.J., & Green, D.P. (2015). When contact changes minds: An experiment on transmission of support for gay equality. *Science, 346*, 6215, 1366-1369. (Retracted on June 5, 2015).

Ladner, G. B. (1976). On Roman attitudes toward barbarians in late Antiquity. *Viator, 7*, 1-26.

Lai, C.K., Marini, M., Lehr, S.A., Cerruti, C., Shin, J.-E. L., & et al. (2014). Reducing implicit racial preferences: I. a comparative investigation of 17 interventions. *Journal of Experimental Psychology, 143*, 4, 1765-1785.

Laland, K., & Brown, G. (2002). *Sense and nonsense: evolutionary perspectives on human behavior.* Oxford, UK: Oxford University Press.

Lasswell, H.D. (1936). *Politics: who gets what, when, how.* New York; London: Whittlesey House, McGraw-Hill.

Lau, R.R., & Sears, D.O. (Eds.) (1986). *Political cognition*. Hillsdale, NJ: Erlbaum.

Lee, V.K., & Harris, L.T. (2014). Dehumanized perception: psychological and neural mechanisms underlying everyday dehumanization. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp.68-87). New York and London: Routledge.

Leidner, B., Castano, E., & Ginges, J. (2013). Dehumanization, Retributive and Restorative Justice, and Aggressive Versus Diplomatic Intergroup Conflict Resolution Strategies. *Personality and Social Psychology Bulletin, 39*, 181-192.

Leyens, J.P. (2014). Humanity forever in medical dehumanization. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp.167-185). New York and London: Routledge.

Leyens, J.P., Rodriguez-Perez, A., Rodriguez-Torres, R., & et al. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology, 31*, 395–411.

Leyens, J.-P., Cortes, B., Demoulin, S., et al. (2003). Emotional prejudice, essentialism, and nationalism: The 2002 Tajfel lecture. *European Journal of Social Psychology, 33*, 6, 703-717.

Lichbach, M.I. (1995). *The rebel's dilemma*. Ann Arbor: University of Michigan Press.

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin, 126*, 109-37.

Lieberman, M.D. (2010). Social cognitive neuroscience. In S.T. Fiske, D.T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology*. Hoboken, NJ: Wiley.

Lieberman, M. D., Schreiber, D., & Ochsner, K. N. (2003). Is Political Cognition Like Riding a Bicycle? How Cognitive Neuroscience Can Inform Research on Political Thinking. *Political Psychology, 24*, 681-704.

Lilienfeld, S. O. (2017). Microaggressions: strong claims, inadequate evidence. *Perspectives on Psychological Science, 12*, 138-169.

Lipset, S. (1960). *Political man*. Garden City, NY: Doubleday.

Liu, H. (2007). New migrants and the revival of overseas Chinese nationalism. *Journal of Contemporary China, 14*, 43, 291-316.

Long, S. L. (Ed.) (1981). *The handbook of political behavior*. Boston, MA: Springer Nature.

Louden, R. B. (2011). *Kant's Human Being*. Oxford, UK: Oxford University Press.

Luban, D. (2015). Human rights pragmatism and human dignity. In R.Cruft, S.M. Liao, & M. Renzo (Eds.), *Philosophical foundations of human rights* (pp.263-278). Oxford, UK: Oxford University Press.

Lukács, G. (1971[1920]). *History and class consciousness: Studies in Marxist dialectics*. Cambridge, Mass.: The MIT Press.

Lutz-Bachmann, M., & Nascimento, A. (Eds.) (2014). *Human rights, human dignity, and cosmopolitan ideals*. Surrey, England; Ashgate.

Lynch, M. (2013). The persistence of Arab anti-Americanism: In the Middle East haters gonna hate. *Foreign Affairs, May/June*, 146-152.

MacKinnon, C.A. (1987). *Feminism unmodified: Discourses on life and law*. Cambridge, Mass.: Harvard University Press.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S.J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: A structural brain imaging analysis. *Hippocampus, 13*, 250–259.

Malabou, C. (2008). *What should we do with our brain?* New York: Fordham University Press.

Marcus, G. E. (2000). Emotions in Politics. *Annual Review of Political Science, 3*, 221-250.

Marcus, G.E. (2013). *Political psychology: neuroscience, genetics, and politics.* New York: Oxford University Press.

Marcus, G.E., Neuman, W.R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. Chicago, IL: Chicago University Press.

Marcuse, H. (1964). *One-dimensional man: studies in the ideology of advanced industrial society*. Boston, MA: Beacon Press.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.

Marshall, T.H. (1965). The right to welfare. *Sociological Review, 13*, 3, 261-272.

Martuzzi, R., Ramani, R., Qiu, M., Rajeevan, N., & Constable, R. T. (2010). Functional connectivity and alterations in baseline brain state in humans. *NeuroImage, 49,* 1, 823-34.

Masters, R.D. (1991). Naturalistic approaches to the concept of justice. *American Behavioral Scientist, 3*, 289-313.

Mathisen, R.W. (2016). Catalogues of barbarians in late Antiquity. In D. Shanzer & R.W. Mathisen (Eds.), *Romans, barbarians, and the transformation of the Roman World* (pp.17-32). Farnham; Burlington: Ashgate.

Mazlish, B. (2001). Civilization in a Historical and Global Perspective. *International Sociology, 16*, 293-300.

Mazziotta, J. C., Toga, A. W., & Frackowiak, R. S. (2000). *Brain Mapping: The Disorders*. San Diego, CL; London: Academic Press.

McCall, C., & Singer, T. (2013). Empathy and the brain. In S. Baron-Cohen, M. Lombardo, & H. Tager-Flusberg (Eds.), *Understanding other minds: perspectives from developmental social neuroscience* (pp. 195-209). Oxford, UK: Oxford University Press.

McCarthy, T. (2009). *Race, empire, and the idea of human development*. Cambridge, UK: Cambridge University Press.

McClelland, J. L., & Ralph, M. A. (2015). *International Encyclopedia of the Social & Behavioral Sciences*. Amsterdam, Netherlands: Elsevier BV.

McCrudden, C. (2008). Human Dignity and Judicial Interpretation of Human Rights. *European Journal of International Law, 19*, 655-724.

McDermott, R. (2004). The Feeling of Rationality: The Meaning of Neuroscientific Advances for Political Science. *Perspectives on Politics, 2*, 691-706.

McDermott, R. (2008). *Presidential leadership: illness and decision-making*. Cambridge, UK: Cambridge University Press.

McDermott, R. (2011). Combining social and biological approaches to political behaviors. *Politics and the Life Sciences, 30*, 98-102.

McDoom, O.S. (2012). The psychology of threat in intergroup conflict: Emotions, rationality, and opportunity in the Rwandan genocide. *International Security*, 37*,* 2, 119-155.

Melleuish, G. (2000). The clash of civilizations: A model of historical development? *Thesis Eleven, 62*, 109-120.

Meloni, M. (2012). On the growing intellectual authority of neuroscience for political and moral theory: sketch for a genealogy. In F. Vander Valk (ed.), *Essays on neuroscience and political theory* (pp. 25-49). London and New York: Routledge.

Merriam, C. E. (1945). *Systematic politics*. Chicago, IL: Chicago University Press.

Mershon, C., & Walsh, D. (2016). Diversity in political science: why it matters and how to get it. *Politics, Groups, and Identities, 4*, 462-466.

Meserve, M. (2009). *Empires of Islam in Renaissance historical thought*. Cambridge, Mass.: Harvard University Press.

Midlarsky, M. I. (1998). Democracy and Islam: Implications for civilizational Conflict and the democratic peace. *International Studies Quarterly, 42*, 485-511.

Mikulincer, M., & Shaver, P. R. (Eds.). (2012). *The social psychology of morality: Exploring the causes of good and evil.* Washington, D.C.: American Psychological Association.

Milgram, S. (1974). *Obedience to authority: an experimental view*. New York: Harper & Row.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences, 7*, 141-144.

Mills, C. (1997), *The racial contract*. Ithaca, NY: Cornell University Press.

Mirandola, P. (1953). *Prospectus for Pico della Mirandola's oration on the dignity of man*. Lexington, KY: Anvil Press.

Mitchell, B.E. (2005). *Black abolitionism: a quest for human dignity*. Maryknoll, NY: Orbis Books.

Mitchell, G. (2010). *The Psychology of Judicial Decision Making*. Oxford, UK: Oxford University Press.

Mitchell, J. P., Neil Macrae, C., & Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: Social-cognitive processing in the medial prefrontal cortex. *NeuroImage, 26*, 251-257.

Mosso, A. (1881). *Über den Kreislauf des Blutes im menschlichen Gehirn* [On blood circulation in the human brain]. Leipzig: Veit.

Mueller, J. (2014). Did history end? Assessing the Fukuyama thesis. *Political Science Quarterly, 129*, 1, 35-54.

Mummendey, A., & Otten, S. (1998). Positive–Negative Asymmetry in Social Discrimination. *European Review of Social Psychology, 9*, 107-143.

Mutua, N. (2016). *Human rights standards: hegemony, law, and politics*. Albany: State University of New York Press.

Mutz, D.C., & Martin, P.S. (2001). Facilitating Communication across Lines of Political Difference: The Role of Mass Media. *The American Political Science Review, 95*(1), 97-114.

Nagel, T. (1974). What is it like to be a bat?. *The Philosophical Review, 83*, 435.

Nakashima, K., Isobe, C., & Ura, M. (2013). How does higher in-group social value lead to positive mental health? An integrated model of in-group identification and support. *Asian Journal of Social Psychology, 16*, 271-278.

National Security Strategy of the United States (2015, February). Issued by the administration of Barack Obama. *National Security Strategy Online Archive*. Retrieved from http://nssarchive.us/wp-content/uploads/2015/02/2015.pdf.

Newman-Norlund, R., Burch, J., & Becofsky, K. (2013). Human Mirror Neuron System (hMNS) Specific Differences in Resting-State Functional Connectivity in Self-Reported Democrats and Republicans: A Pilot Study. *JBBS, 03*, 341-349.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Nisbett, R. E. (2003). *Geography of thought: how Asians and Westerners think differently – and why.* New York: Free Press.

Nittono, H., Kawamoto, T., & Ura, M. (2015). Intrapersonal and interpersonal processes of social exclusion. *Frontiers in Neuroscience, 9*, 62.

Novak, M. (1999). Human dignity, human rights. *First things: A Monthly Journal of Religion and Public Life, 97,* 39-42.

Nussbaum, M, & Cohen, J. (2002a). *For love of country?* Boston: Beacon Press.

Nussbaum, M. (2002b). *Gender Justice, Development, and Rights*. None: Oxford University Press (OUP).

O'Hagan, J. (2007). Discourses of civilizational identity. In In M. Hall & P.T. Jackson (Eds.), *Civilizational identity: The production and reproduction of "civilizations" in International relations* (pp.15-32). Basingstoke; New York: Palgrave Macmillan.

O'Shea, M. (2005). *The brain: A very short introduction*. Oxford, UK: Oxford University Press.

Ochsner, K. N. (2007). Social cognitive neuroscience: Historical development, core principles, and future promise. In A. Kruglanksi & E.T. Higgins (Eds.), *Social Psychology: A Handbook of Basic Principles* (2nd ed., pp. 39 – 66). New York: Guilford Press.

Ochsner, K. N., & Schacter, D. L. (2000). A social cognitive neuroscience approach to emotion and memory. In J. C. Borod (Ed.), *The neuropsychology of emotion* (pp. 163-193). New York: Oxford University Press.

Olivola, C.Y., & Sussman, A.B. (2014). Many behavioral tendencies associated with right-leaning (conservative) political ideologies are malleable and unrelated to negativity. *Behavioral and Brain Sciences, 37*, 3, 323-324.

Opotow, S. (1990). Moral Exclusion and Injustice: An Introduction. *Journal of Social Issues, 46*, 1, 1-20.

Oyserman, D., & Lee, S. W. S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism.. *Psychological Bulletin, 134*, 311-342.

Pagel, M.D. (2012). *Wired for culture: origins of the human social mind*. New York: W.W. Norton.

Palmer, J., & Smith, D. (2010). Promoting Human Dignity: An evaluation of a programme for racially motivated offenders. *Probation Journal, 57*, 368-382.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions.* New York: Oxford University Press.

Pardo, M. S., & Patterson, D. (2013). *Minds, brains, and law: the conceptual foundations of law and neuroscience*. New York: Oxford University Press.

Parekh, B. (2000). *Rethinking multiculturalism: cultural diversity and political theory.* Cambridge, Mass.: Harvard University Press.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology, 60*, 173-196.

Patömaki, H., & Wight, C. (2000). After postpositivism? The promises of critical realism. *International Studies Quarterly, 44*, 2, 213-237.

Patten, A. (2014). *Equal recognition: The moral foundations of minority rights*. Princeton, NJ: Princeton University Press.

Patton, T. O. (2004). Reflections of a black woman professor: Racism and sexism in academia. *Howard Journal of Communications, 15*, 185-200.

Penelhum, T. (1970). *Survival and disembodies existence*. London: Routledge.

Penrose, R. (1990). *The Emperor's new mind*. Oxford, UK: Oxford University Press.

Pentassuglia, G. (2012). The Strasbourg Court and Minority Groups: Shooting in the Dark or a New Interpretive Ethos?, *19*, 1-23.

Pereira, C., Vala, J., & Leyens, J.-P. (2009). From infra-humanization to discrimination: The mediation of symbolic threat needs egalitarian norms. *Journal of Experimental Social Psychology, 45*, 2, 336-344.

Peroni, L. (2015) Minorities before the European Court of human rights: Democratic pluralism unfolded. In J. Boulden & W. Kymlicka (Eds.), *International approaches to governing ethnic diversity* (pp.25-50). Oxford, UK: Oxford University Press.

Perrett, D.I., Rolls, E.T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47*, 329-342.

Petersen, R. (2011). *Western intervention in the Balkans: the strategic use of emotion in conflict*. Cambridge, UK: Cambridge University Press.

Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin, 23*, 173-185.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.

Piccolino, M. (2013). *Shocking frogs: Galvani, Volta, and the electric origins of neuroscience*. Oxford, UK: Oxford University Press.

Platek, S. M., Krill, A. L., & Wilson, B. (2009). Implicit trustworthiness ratings of self-resembling faces activate brain centers involved in reward. *Neuropsychologia, 47*, 289-293.

Platek, S. M., & Krill, A. L. (2009). Self-face resemblance attenuates other-race face effect in the amygdala. *Brain Research, 1284*, 156-160.

Polakow-Suransky, S. (2016, November 1). The ruthlessly effective rebranding of Europe's new far right. *The Guardian Newspaper*. Retrieved from https://www.theguardian.com/world/2016/nov/01/the-ruthlessly-effective-rebranding-of-europes-new-far-right.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*, 59-63.

Poldrack, R. A. (2008). The role of fMRI in Cognitive Neuroscience: where do we stand? *Current Opinion in Neurobiology, 18*, 223-227.

Pornpattananangkul, N., Cheon, B. K., & Chiao, J. Y. (2014). The role of negativity bias in political judgment: A cultural neuroscience perspective. *Behavioral and Brain Sciences, 37*, 3, 325-6.

Power, S. (2002). *A problem from hell: America and the age of genocide*. New York, NY: Basic Books.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral Brain Sciences, 1*, 515-526.

Putnam, R.D. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.

Qian, S. (1971). *Records of the grand historian of China*. New York: Columbia University Press.

Rachlin, H. (1994). From overt behavior to hypothetical behavior to memory: Inference in the wrong direction. *Behavioral Brain Sciences, 17*, 147.

Raichle, M. E., & Snyder, A.Z. (2000). A default mode of brain function: A brief history of an evolving idea. *NeuroImage, 37*, 4, 1083-1090.

Ramachandran, V. (2011). *The tell-tale brain: A neuroscientist's quest for what makes us human*. New York: Norton.

Rankin, R. E., & Campbell, D. T. (1955). Galvanic skin response to Negro and white experimenters. *The Journal of Abnormal and Social Psychology, 51*, 30-33.

Rawls, J. (1971). *A theory of justice*. Cambridge, Mass.: Belknap Press.

Rawls, J. (1993). *Political liberalism*. New York: Columbia University Press.

Raz, J. (1994). *Ethics in the public domain: essays in the morality of law and politics*. Oxford, UK: Clarendon Press

Reddy, W.M. (2001). *The Navigation of Feeling: A Framework for the History of Emotions.* New York: Cambridge University Press.

Redlawsk, D. P. (Ed.). (2006). *Feeling politics: emotion in political information processing*. Basingstoke, England: Palgrave Macmillan.

Reich, R. (2005). Minors within minorities: A problem for liberal multiculturalists. In A. Eisenberg & J. Spinner-Halev, *Minorities within minorities: equality, rights, and diversity* (pp.209-226). Cambridge, UK: Cambridge University Press.

Richards, R. (1987). *Darwin and the emergence of evolutionary theories of mind and*

behavior. Chicago, IL: University of Chicago Press.

Rippon, G. Jordan-Young, R., Kaiser, A., & Fine, C. (2014) Recommendations for sex/gender neuroimaging research: Key principles and implications for research design, analysis and interpretation. *Frontiers in Human Neuroscience*, *8*, 650

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169-192.

Robinson, H. (1982). *Matter and sense*. Cambridge, UK: Cambridge University Press.

Robinson, J. (1965). *And the crooked shall be made straight: The Eichmann Trial, the Jewish catastrophe and Hannah Arendt's narrative*. New York and London: The Macmillan Company.

Roccas, S., Klar, Y., & Liviatan, I. (2006). The paradox of group-based guilt: Modes of national identification, conflict vehemence, and reactions to the in-group's moral violations. *Journal of Personality and Social Psychology, 91*, 4, 698-711.

Roland, P. (1993). *Brain activation*. New York: Wiley-Liss.

Rolls, E.T. (2012). *Neuroculture: On the implications of brain science*. Oxford, UK: Oxford University Press.

Rolls, E. T., & Deco, G. (2010). *The noisy brain stochastic dynamics as a principle of brain function*. Oxford, UK: Oxford University Press.

Ronquillo J., Denson T. F., Lickel B., Lu Z.-L., Nandy A., & Maddox K. B. (2007). The effects of skin tone on race-related amygdala activity: An fMRI investigation. *Social Cognitive and Affective Neuroscience, 2*, 39-44.

Rosen, M. (1996). *On voluntary servitude: false consciousness and the theory of ideology*. Cambridge, UK: Polity Press.

Rosen, M. (2012). *Dignity: its history and meaning*. Cambridge, Mass.: Harvard University Press.

Röska-Hardy, L. S., & Neumann-Held, E.M., (2009). *Learning from animals?: Examining the nature of human uniqueness*. Hove; New York: Psychology Press.

Roy, J.-M. (2016). From intersubjectivity to International Relations. In Y. Ariffin, J.-M. Coicaud, & V. Popovski (Eds.), *Emotions in International politics* (pp.65-79). Cambridge, UK: Cambridge University Press.

Said, E. (2001, October 4). The clash of ignorance. *The Nation*. Retrieved from https://www.thenation.com/article/clash-ignorance/.

Said, E. (2004). *From Oslo to Iraq and the road map*. New York: Pantheon Books.

Said, E. (1978). *Orientalism*. New York: Pantheon Books.

Said, E. (1994). *Culture and imperialism*. New York: Vintage Books.

Sandel, M. (1982). *Liberalism and the limits of justice*. Cambridge, UK: Cambridge University Press.

Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive

neuroscience: Toward strong inference in attributing function to structure. *American Psychologist, 51*, 13-21.

Sartre, J.P. (1948). *Anti-Semite and Jew*. New York: Schocken Books.

Saurin, J. (2006). International relations as the imperial illusion; or, the need to decolonize IR. In B.G. Jones (Ed.), *Decolonizing International Relations* (pp.23-42). Lanham: Rowman & Littlefield.

Savulescu, J., & Kahane, G. (2011). Disability: a welfarist approach. *Clinical Ethics, 6*, 45-51.

Savulescu, J. (2006). End-of-life decisions. *The Foundation Years, 2*, 151-154.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *NeuroImage, 19*, 1835-1842.

Scanlon, T. M. (1973). Rawls' Theory of Justice. *University of Pennsylvania Law Review, 121*, 1020-1069.

Schmitz, S.S., & Höppner, G. (2014). Neurofeminism and feminist neurosciences: a critical review of contemporary brain research. *Frontiers in human neuroscience, 8*, 546.

Schutt, R. K., Seidman, L. J., & Keshavan, M. S. (2015). *Social Neuroscience*. Cambridge, Mss.: Harvard University Press.

Schwartz, S. H., Caprara, G. V., & Vecchione, M. (2010). Basic Personal Values, Core Political Values, and Voting: A Longitudinal Analysis. *Political Psychology, 31*, 3, 421-452.

Selway, J. S. (2011). The measurement of cross-cutting cleavages and other multidimensional cleavage structures. *Political Analysis, 19*, 1, 48-65.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology, 54*, 558-568.

Sen, A. (1999). *Development as freedom*. New York: Knopf.

Sered, S. (2014, November 17). Why brain science won't cure poverty. *The Conversation U.S.*. Retrieved from http://theconversation.com/why-brain-science-wont-cure-poverty-33920.

Shapin, S., & Schaffer, S. (1985). *Hobbes and the air-pump: Hobbes, Boyle and the experimental life.* Princeton, NJ: Princeton University Press.

Shelby, T. (2005). *We who are dark: the philosophical foundations of Black solidarity*. Cambridge, Mass.: Belknap Press.

Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *NeuroImage, 61*, 786-797.

Sheng, F., Han, X., & Han, S. (2015). Dissociated neural representations of pain expressions of different races. *Cerebral Cortex 26*, 3, 1221-1233.

Sherif, M., Harvey, O.J., White, J.B., et al. (1961). *Intergroup conflict and cooperation: The Robbers cave experiment*. Norman, OK: University of Oklahoma Press.

Shook, N. J., & Fazio, R. H. (2009). Political ideology, exploration of novel stimuli, and attitude formation. *Journal of Experimental Social Psychology, 45*, 995-998.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge; New York: Cambridge University Press.

Simmons, J. (2015). Human rights, natural rights, and human dignity. In R.Cruft, S.M. Liao, & M. Renzo (Eds.), *Philosophical foundations of human rights* (pp.138-153). Oxford, UK: Oxford University Press.

Simon, (1982). *Models of bounded rationality*. Cambridge, Mass.: The MIT Press.

Simon, H. A. (1985). Human nature in politics: The dialogue of psychology with political science. *American Political Science Review, 79*, 293-304.

Singer, P. (1979). *Practical ethics*. Cambridge, UK: Cambridge University Press.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology, 24*, R875-R878.

Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist, 8*, 69-78.

Slomp, G. (2000). *Thomas Hobbes and the political philosophy of glory.* New York: St. Martin's Press.

So, J.F., & Bunker, E.C. (1995). *Traders and raiders on China's Northern frontier*. Seattle: University of Washington Press.

Staub, E. (1989). *The roots of evil: the origins of genocide and other group violence*. Cambridge, UK: Cambridge University Press.

Staub, E. (2010). *Overcoming Evil*. Oxford, UK: Oxford University Press.

Stein, E. (1989[1917]). *On the problem of empathy*, transl. Waltraut Stein. Washington, DC: ICS Publications.

Strauss, L. (1953). *Natural right and history*. Chicago, IL: University of Chicago Press.

Sue, D. W. (2009). Racial microaggressions and worldviews. *American Psychologist, 64*, 220-221.

Sue, D. W. (2017). Microaggressions and "Evidence" *Perspectives on Psychological Science, 12*, 170-172.

Swencionis, J. K., & Fiske, S. T. (2014). More human: individuation in the 21st century. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp.276-293). New York and London: Routledge.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology. 33*, 1-39.

Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology.* Cambridge, UK: Cambridge University Press.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7-24). Chicago: Nelson-Hall.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 5, 96-103.

Talmi, D., Hurlemann, R., Patin, A., & Dolan, R.J. (2010). Framing effect following bilateral amygdala lesion. *Neuropsychologia, 48*, 6, 1823-1827.

Tarrow, (1996). Making social science work across space and time: A critical reflection on Robert Putnam's Making Democracy Work. *American Political Science Review, 90*, 2, 389-397.

Tasioulas, J. (2012). Towards a philosophy of human rights. *Current Legal Problems, 65*, 1-30.

Tatum, D.C. (2010). *Genocide at the dawn of the twenty-first century: Rwanda, Bosnia, Kosovo and Darfur*. New York, NY: Palgrave Macmillan.

Taylor, C. (1989). *Sources of the self: The making of the modern identity*. Cambridge, Mass.: Harvard University Press.

Taylor, C. (1994). *Multiculturalism and the "politics of recognition": an essay*. Princeton, NJ: Princeton University Press.

Taylor, K.-Y. (2016). *From #BlackLivesMatter to Black liberation*. Chicago, IL: Haymarket Books.

Terbeck, S., Kahane, G., McTavish, S. et al. (2012). Propranolol reduces implicit negative racial bias. *Psychopharmacology, 222*, 419.

Terbeck, S., Kahane, G., McTavish, S. et al. (2015). β-Adrenoceptor blockade modulates fusiform gyrus activity to black versus white faces, *Psychopharmacology, 232*, 2951.

Terbeck, S. (2016). *The social neuroscience of intergroup relations: Prejudice, can we cure it?*. Cham, Switzerland: Springer International Publishing.

Terry, B. (2013). Critical race theory and the tasks of political philosophy: On Rawls and the racial contract. *Conference Paper*. National Conference of Black Political Scientists, March 2013, Wilmington, D.E.

Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social Psychology, 45*, 118-126.

Theodoridis, A. G., & Nelson, A. J. (2012). Of BOLD Claims and Excessive Fears: A Call for Caution and Patience Regarding Political Neuroscience. *Political Psychology, 33*, 27-43.

Thiele, L.P. (2006). *The heart of judgment: Practical wisdom, neuroscience, and narrative*. Cambridge, UK: Cambridge University Press.

Tingley, D. (2006). Neurological imaging as evidence in political science: a review, critique, and guiding assessment. *Social Science Information, 45*, 5-33.

Todorov, A., Harris, L. T., & Fiske, S. T. (2006). Toward socially inspired social neuroscience. *Brain Research, 1079*, 76-85.

Todorov, A., Fiske, S. T., & Harris, L. T. (2011). *Social Neuroscience*. Oxford, UK: Oxford University Press.

Todorov, A., Fiske, S.T., & Prentice, D.A. (2011). *Social neuroscience: Toward understanding the underpinning of the social mind*. Oxford, New York: Oxford University Press.

Todorova, M. (1997). *Imagining the Balkans*. Oxford, UK: Oxford University Press.

Tolan, J.V. (2002). *Saracens: Islam in the Medieval European Imagination*. New York: Columbia University Press.

Tolan, J., Laurens, H., & Veinstein, G. (2012). *Europe and the Islamic world: A history*. Princeton, NJ: Princeton University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Science, 7*, 153-156.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*, 859-883.

Trix, F. (2014). Peace-mongering in 1913: The Carnegie International commission of inquiry and its report on the Balkan wars. *First World War Studies, 5*, 2, 147-162.

Tuck, R. (1991). Introduction. In T. Hobbes, *Leviathan* (pp.ix-xlv). Cambridge, UK: Cambridge University Press.

Tusche, A., Kahnt, T., Wisniewski, D., & Haynes, J. (2013). Automatic processing of political preferences in the human brain. *Neuroimage, 72*, 174–182.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*, 1124-1131.

U.S. Census Bureau. Overview of race and Hispanic origin: 2010. Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf.

Uttal, D. H. (2001). Making sense of the development of spatial cognition. *Trends in Cognitive Sciences, 5*, 316-317.

Utz, S. (2004). Self-Construal and Cooperation: Is the Interdependent Self More Cooperative Than the Independent Self?. *Self and Identity, 3*, 177-190.

Vaes, J., & Paladino, M.P. (2010). The uniquely human content of stereotypes. *Group Processes and Intergroup Relations, 13*, 23-39.

Vaes, J., Paladino, M. P., & Leyens, J. (2006). Priming uniquely human emotions and the in-group (but not the out-group) activates humanity concepts. *Eur. J. Soc. Psychol., 36*, 169-181.

Vaes, J., Paladino, M. P., & Magagnotti, C. (2011). The human message in politics: The impact of emotional slogans on subtle conformity. *The Journal of Social Psychology, 151*, 162-179.

Vaes, J., Paladino, M.P., Castelli, L., Leyens, J.-P., & Giovanazzi, A. (2003). On the behavioral consequences of infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *Journal of Personality and Social Psychology, 85*, 6, 1016-1034.

Vaish, A., & Warneken, F. (2012) Social-cognitive contributors to young children's empathic and prosocial behavior. In J. Decety (Ed.), *Empathy: from bench to bedside*. Cambridge, Mass.: The MIT Press.

Vander Valk, F. (Ed.) (2012). *Essays on neuroscience and political theory: Thinking the body politic*. London; New York: Routledge.

Vanderbilt, D., Young, R., MacDonald, H. Z., Grant-Knight, W., Saxe, G., & Zuckerman, B. (2008). Asthma Severity and PTSD Symptoms Among Inner City Children: A Pilot Study. *Journal of Trauma & Dissociation, 9*, 191-207.

Vasiljevic, M., & Viki, G.T. (2014). Dehumanization, moral punishment, and public attitudes to crime and punishment. In P.G. Bain, J. Vaes, & J.-P. Leyens (Eds.), *Humanness and dehumanization* (pp.129-146). New York and London: Routledge.

Vauclair, J., & Perret, P. (2003). The cognitive revolution in Europe: Taking the developmental perspective seriously. *Trends in Cognitive Sciences, 7*, 7, 284-285.

Vaughan, G. M., Tajfel, H., & Williams, J. (1981). Bias in reward allocation in an intergroup and an interpersonal context. *Social Psychology Quarterly, 44*, 37.

Vidulich, R.N., & Krevanick, F.W. (1966). Racial attitudes and emotional responses to visual representations of the Negro. *Journal of Social Psychology, 68*, 1, 85-93.

Viki, G. T., Osgood, D., & Phillips, S. (2013). Dehumanization and self-reported proclivity to torture prisoners of war. *Journal of Experimental Social Psychology, 49*, 3, 325-328.

Vlassopoulos, K. (2013). *Greeks and barbarians*. Cambridge; New York: Cambridge University Press.

Voegelin, E. (2000[1952]). *Published Essays: 1940-1952*. Columbia, MO: University of Missouri Press.

Vorauer, J.D., Main, K.J., & O'Connell, G.B. (1998). How do individuals expect to be viewed by members of lower status groups? Content and implications of meta-stereotypes. *Journal of Personality and Social Psychology, 75*, 4, 917-937.

Vuillamy, E. (1998). Bosnia: The crime of appeasement. *International Affairs, 74*, 1, 73-91.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4*, 274 – 290.

Wagner, D.D. (2015). Mentalizing. *Brain mapping: An encyclopedic reference* (pp.143-146). Amsterdam, Netherlands: Academic Press.

Wahl, R. (2013). Policing, values, and violence: Human rights education with law enforcers in India. *Journal of Human Rights Practice, 5*, 2, 220-242.

Waldron, J. (2000). *Citizenship in diverse societies*. Oxford; New York: Oxford University Press.

Waldron, J. (2003). Toleration and reasonableness. In C. McKinnon & D. Castiglione, *The culture of toleration in diverse societies: Reasonable tolerance* (pp. 13-37). Manchester, UK: Manchester University Press.

Waldron, J. (2012). *Dignity, rank, and rights*. Oxford; New York: Oxford University Press.

Walzer, M. (2004). *Politics and passion: Toward a more egalitarian liberalism*. New Haven, Conn.: Yale University Press.

Ward, C., Fox, S., Wilson, J., Stuart, J., & Kus, L. (2010). Contextual influences on acculturation processes: The roles of family, community and society. *Psychological Studies, 55*, 26-34.

Warnick, J. E., & Landis, D. (2015). Introduction. In J.E. Warnick & D. Landis (Eds.), *Neuroscience in intercultural contexts* (pp.1-30). New York: Springer New York.

Watson, J. (1913). Psychology as a behaviorist views it. *Psychological Review, 20*, 158-177.

Waytz, A., Hoffman, K.M., & Trawalter, S. (2013). A superhumanization bias in Whites' perceptions of Blacks. *Social Psychological and Personality Science, 6*, 3, 352-359.

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness–loyalty tradeoff. *Journal of Experimental Social Psychology, 49*, 1027-1033.

Wendt, A. E. (1987). The agent-structure problem in international relations theory. *International Organization, 41*, 335.

West, R. (1941). *Black lamb and grey falcon: A journey through Yugoslavia.* New York: Viking Press.

Westen, D. (2008). *The political brain: the role of emotion in deciding the fate of the nation*. New York: Public Affairs.

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election. *Journal of Cognitive Neuroscience, 18*, 1947-1958.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice. *Psychological Science, 16*, 56-63.

Wiegele, T.C. (1979). *Biopolitics: search for a more human political science*. Boulder, CO: Westview Press.

Wiegele, T.C. (1985). *Leaders under stress: a psychophysiological analysis of international crises.* Durham, NC: Duke University Press.

Wight, C. (2006). *Agents, structures, and International relations: Politics as ontology*. Cambridge, UK: Cambridge University Press.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms

exposure to a face. *Psychological Science, 17*, 592-598.

Wilson, E.O. (1975). *Sociobiology: the new synthesis*. Cambridge, Mass.: Belknap Press.

Wilson, F.R. (1998). *The hand: How its use shapes the brain, language, and human culture*. New York: Pantheon Books.

Winter, D. G. (2000). Power, Sex, and Violence: A Psychological Reconstruction of the 20th Century and an Intellectual Agenda for Political Psychology. *Political Psychology, 21*, 383-404.

Winter, B. (2008). *Hijab and the Republic: Uncovering the French headscarf debate*. Syracuse, NY: Syracuse University Press.

Wolff, M. (2016, November 18). Ringside with Steve Bannon at Trump Tower as the President-elect's strategist plots "an entirely new political movement" (Exclusive). *The Hollywood Reporter*. Retrieved from http://www.hollywoodreporter.com/news/steve-bannon-trump-tower-interview-trumps-strategist-plots-new-political-movement-948747.

Wong, R.Y., & Wong, Y.Y. (2005). Dynamic influences of culture on cooperation in the prisoner's dilemma. *Psychological Science, 16*, 6, 429-434.

Wu, F.H. (2002). *Yellow: Race in America beyond black and white*. New York: Basic Books.

Wu, Y., Wang, C., He, X., Mao, L., & Zhang, L. (2010). Religious beliefs influence neural substrates of self-reflection in Tibetans. *Soc Cogn Affect Neurosci, 5*, 324-331.

Yang, J. (2008, May, 5). The facets of Chinese nationalism. *The Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2008/05/04/AR2008050401599.html.

Young, A.W., Rowland, D., Calder, A.J., Etcoff, N.L., Seth, A., & Perrett, D.I. (1997) Facial expression megamix: Tests of dimensional and category accounts of emotion recognition, *Cognition, 63*, 3, 271-313.

Young, E. (2009) Why we're liberal, why we're conservative: A cognitive theory on the origins of ideological thinking. *Unpublished doctoral dissertation*. Department of Political Science: Stony Brook University.

Young, L., & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *NeuroImage, 40*, 1912-1920.

Young, L., & Waytz, A. (2013). Mind attribution is for morality. In S. Baron-Cohen, M. Lombardo, & H. Tager-Flusberg (Eds.), *Understanding other minds: perspectives from developmental social neuroscience* (pp. 93-101). Oxford, UK: Oxford University Press.

Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology, 40*, 560-567.

Zakaria, F. (2001, October 14). The politics of rage: Why do they hate us?. *Newsweek*.

Retrieved from http://www.newsweek.com/politics-rage-why-do-they-hate-us-154345.

Zakaria, F. (1994). Culture is destiny: A conversation with Lee Kuan Yew. *Foreign Affairs, 73*, 2, 109-126.

Zamboni, G., Gozzi, M., Krueger, F., Duhamel, J., Sirigu, A., & Grafman, J. (2009). Individualism, conservatism, and radicalism as criteria for processing political beliefs: A parametric fMRI study. *Social Neuroscience, 4*, 367–383.

Zhang, L., Tiangang, Z., Zhang, J., Liu, Z., Fan, J., & Ying, Z. (2006). In search of the Chinese self: An fMRI study. *Science in China, 49*, 89 – 96.

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage, 34*, 1310-1316.

Zimmerman, W. (1996). *Origins of a catastrophe: Yugoslavia and its destroyers – America's last ambassador tells what happened and why.* New York: Times Books.

Zuo, X., & Han, S. (2013). Cultural experiences reduce racial bias in neural responses to others' suffering. *Cult. Brain, 1*, 34-46.

Zuriff, G. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.