



Hate speech, volition, and neurology

Alexandra B. Roginsky^{1,†} and Alexander Tsesis^{2,*‡}

1. Department of Surgery, Centegra Health System, McHenry, Illinois, USA

2. Loyola University Chicago School of Law, Chicago, USA

*Corresponding author. E-mail: atsesis@luc.edu

In ‘A Hypothetical Neurological Association between Dehumanization and Human Rights Abuse’,¹ Gail Murrow and Richard Murrow posit a biological explanation of how hate speech can spur violence, not only among individuals but, even, on a societal scale. They elaborate historical examples, cite to neuronal studies on patterns of responses in observation of pain and suffering to explain the dehumanization that often results from hate propaganda. The authors’ points about the harmful effects of hate speech are salient and erudite, and we commend them for their deeply thought out work. However, their scientific premises rely on significant extrapolations from neurological theories on manifold human behavior.

Hate speakers rely on dehumanizing images to justify exclusion, discrimination, and, in genocidal cases, elimination of identifiable groups. Dehumanization can be both an attack on the target’s dignity and a justification for harmful actions. Statements dehumanizing hated groups often influence the commission of discriminatory conduct. The critical role of rhetoric in motivating nefarious action is evident in the histories of genocides in Germany, Turkey, Sudan, and Rwanda. In all of these countries, the official spread of malignant and distorted images of the other (Jews, Armenians, Darfuris, and Tutsis, respectively) made it easy to bring the hated groups into disrepute with the population and cleared the way to their mass killing and divestment of property.

That neurology can help explain these phenomena cannot be doubted, given that our brains are intrinsic to thought, but that neurology is a sufficient explanation of dehumanization is suspect given the socialization of repeated group defamations. Chiefly, Murrow and Murrow rely on the scientific discussions on ‘mirror neurons’, a relatively

[†] Alexandra B. Roginsky, MD, FACS, General and Robotic Surgeon at Centegra Health System, McHenry, Illinois.

[‡] Alexander Tsesis, JD, Professor of Law, Loyola University Chicago School of Law.

¹ Gail B. Murrow and Richard Murrow, *A Hypothetical Neurological Association between Dehumanization and Human Rights Abuses*, *J. Law & BIOSCI.* (2015), available at <http://jlb.oxfordjournals.org/content/early/2015/06/08/jlb.lsv015.full> (accessed 12 December 2015).

new concept in neurobiology, to explain a neurological link between dehumanization and violence on a societal scale.

It is important to understand the original scientific literature on mirror neurons to better assess whether, and how, these neuronal processes might be involved in complex social and class behaviors. In 1992, Giuseppe di Pellegrino et al., published the first scientific paper on mirror neurons, a completely novel concept at the time. In ‘Understanding Motor Events, a Neurophysiological Study’, di Pellegrino et al. describe neuronal activity in the ventral premotor region of the macaque monkey both when a specific motor activity is performed, and observed.² The original publication described very detailed physical hand movements (grasping, lifting, etc.) and the accompanying neuronal activities. The same research group coined the specific term, ‘mirror neurons’, in a subsequent publication.³ There have been many publications on the subject in subsequent years, both in scientific peer review and lay journals, with a variety of results, as expected.

Before considering the theory’s applicability to hate speech, it is important to carefully examine the primary scientific literature on mirror neurons. The studies are very elegant and highly specific in describing neurobiology. In fact, the definition of mirror neurons, by the original authors, states that they are neurons that ‘discharged both during monkey’s active movements and when the monkey observed meaningful hand movements made by the experimenter’.⁴ This is a highly technical explanation based on observation and movement. From a cellular perspective, mirror neurons have modulated activity both by ‘action execution and action observation’, in contrast with sensory and motor neurons, whose firing is tied to one of the above functions, but not to both.⁵

The recording of neuronal firing is very far removed from descriptions of genocidal hatred, and the complex societal norms that encouraged it. These societal phenomena often incorporate propaganda to influence others, especially those who are most impressionable. We believe that explaining the effects of hate speech on society in terms of simple execution and observation oversimplifies indoctrination, internalization, and affectation.

The mirror neuron studies often involved training monkeys for months to imitate specific hand motions, and then studied neuronal discharge of specific areas and neurons after implantation of titanium electrodes.⁶ Such studies have not been, and ethically cannot be, conducted in humans. To make scientific statements, one must be able to demonstrate reproducibility. Therefore, investigators rely on a different mechanism to investigate neuronal activities, mostly through imaging technologies, such as functional magnetic resonance imaging (fMRI).⁷

² Giuseppe di Pellegrino et al., *Understanding Motor Events, a Neurophysiological Study*, 91 EXP. BRAIN RES. 176 (1992).

³ Vittorio Gallese et al., *Action Recognition in the Premotor Cortex*, 119 BRAIN 593 (1996).

⁴ *Id.* at 4.

⁵ James M. Kilner and Roger N. Lemon, *What We Know Currently about Mirror Neurons*, 23 CURR. BIOL. 1057 (2013).

⁶ Luca Bonini et al., *Ventral Premotor and Inferior Parietal Cortices Make Distinct Contribution to Action Organization and Intention Understanding*, 20 CEREB. CORTEX 1372, 1385 (2010).

⁷ *Id.* at 6.

The difference in study format necessitates a significant extrapolation of hypotheses about mirror neuron functions between species. In fact, mirror neuron function studies on humans report mixed results.⁸ However, in this paper, it is not our intent to argue that no mirror neurons exist in humans. Rather, our aim is to draw attention to the fact that mirror neurons have been studied by scientists in order to explain physical movements and observations associated with motor learning, a primitive function, not highly complex, societal events like animus, advocacy to violence, and other forms of discrimination.

At every step in the tragic process of hateful instigation people have the ability to reflect on the hateful rhetoric and acknowledge the humanity of the other. Thus Hannah Arendt, whose work Murrow and Murrow draw upon for the supposed 'unthinking component' of Nazi functionaries, was wrong to characterize the organizer of European Jewish deportation to concentration camps, Adolf Eichmann, as an unreflective functionary engaged in automatic conduct.⁹ As a recent study of Eichmann's post-War writings reveals, he was a fanatical ideologue whose participation in genocide was motivated by a murderous loathing of Jews.¹⁰ He knowingly acted on the gross belief, one that he shared with other Nazis, that Jews were the chief enemies of mankind who needed to be destroyed like vermin. In analysing his actions, what is important is not a motor neuron-centered explanation of reflexive behaviors but how his actions reflected, were influenced by, and in turn swayed the Holocaust. Arendt sought to build a picture of a thoughtless automaton in a political machine; in reality, Eichmann, as any other perpetrator of genocide, was a thoughtful, calculating, consciously volitional functionary, acting on destructive beliefs about the indelible character of the other, and through his hatred performed in a highly organized manner with nothing left to chance.

Much as we agree with Murrow and Murrow's depiction of hate speech as a dehumanizing force that blunts human emotions about the hated others, we are concerned about their claim that neural mechanisms sufficiently explain the blunting of empathy because this inadvertently exonerates the perpetrators. Group hatred is not merely a neuron-transmitted sensation of animus but breakdown of empathy, human concern, and sense of fairness. The object of group ire is regarded not only as an animal without political rights, artistic creativity, and self-government. Rather, dehumanizing hate speech conceives outgroup members to be naturally, civically, or religiously nefarious and therefore justifiably excluded from the benefits reserved for persons with favored characteristics. Thus, hatred just as empathy appeals to the rational faculty, giving reason for action and identification. The complexity of this process is inexplicable as a simple matter of neural pain signal transmission. Genocide and the large-scale murder require active participation and persistence. Interestingly, in the mirror neuron studies, there was a large variety of rate in firing of the specified neurons, even when consistent methods were used.¹¹

Socialization is essential to the internalization of aversion and hatred. Acquisition of misethnic attitudes begins early in life through childhood observation, language development, formal education, and behavioral and verbal imitation. The influence of hate

⁸ *Id.*

⁹ HANNAH ARENDT, *EICHMANN IN JERUSALEM* (1963).

¹⁰ BETTINA STANGNETH, *EICHMANN BEFORE JERUSALEM: THE UNEXAMINED LIFE OF A MASS MURDERER* (2014).

¹¹ Kilner & Lamon, *supra* note 7, at 5.

speech relies on the ability of stereotypes to be readily identified by ordinary persons. Saying that Jews kill Christian children, for instance, causes not only a neurological response but also draws upon a historic trope of Jews as dangerous interlopers. Likewise, when a bigot speaks of Latinos as being lazy outsiders, gypsies as thieves, and gays as pedophiles, he seeks to elicit negative audience response that relies not only on mechanistic mirroring but on the speaker's ability to invoke widespread and deeply held convictions.

Group defamations convey widespread prejudices through juxtapositions, distinctions, and hyperboles. Discriminatory definitions glorify ingroup character, morality, fidelity, intelligence, beauty, and devalue the intrinsic dignity of outgroups. They render the world into distinct camps with impenetrable boundaries or race, ethnicity, sex, sexual orientation, religion, or nationality. Negative stereotypes, therefore, elicit illusions of grandeur. Ingroups describe themselves in positive terms. Outgroups, on the other hand, are perceived as dangerous to culture and institutions.

Derogatory generalizations about minorities direct cultural thought about the hated others that engage far more than mirrored automatic responses. They are like blinders that restrain the range of public perceptions of outgroups to a narrow set of defamations, charged with cultural meanings, that misrepresent outgroup characteristics. Eventually, stereotypes become clichés and supposed truisms. Derogatory judgments about an outgroup evolve into social definitions that are applied to entire group.

In summary, societal hatred is an extremely complex social phenomenon that takes years to form and evolve, and subsequently, lead to action. Propaganda tools include visual cartoons, 'catchy' and memorable phrases ('the Jews are our misfortune', 'He gyped me', 'Yellow Peril', and the like), and overt or implicit incitements to violence or discrimination. Murrow and Murrow provide an extremely articulate discussion about the poisonous nature of hate speech. However, connecting the scientific concept of mirror neurons as a potential explanation for individuals' capacities to participate in discrimination, and worse, is too much of an extrapolation. The original authors of the scientific papers on mirror neurons described them as part of a motor function process, not a theory to help explain the darkest times of history and human behavior.