Neuroscience cannot answer these questions: a response to G. and R. Murrow’s essay hypothesizing a link between dehumanization, human rights abuses and public policy

Morris B. Hoffman*

District Court, Second Judicial District, Denver, Colorado, USA
Corresponding author. E-mail: morris.hoffman@judicial.state.co.us

ABSTRACT

The Murrows’ paper, ‘A hypothetical link between dehumanization and human rights abuses’, in which they propose that neuroscience may answer some difficult public policy questions, including questions about the First Amendment, is an unfortunate foray into law and public policy unjustified by the current state of neuroscience. Neuroscientific insights may one day have important implications for the law, and for some of the folk psychological assumptions embedded in the law, but they will never change the words of the written Constitution, or answer difficult policy questions in the interstices of those words. Suggesting that neuroscience can today inform these questions does a disservice to science, law and the complexity of the human condition.

KEYWORDS: constitutional law, empathy, First Amendment, mirror neurons, neuroscience, policy

* Morris B. Hoffman, District Judge, Second Judicial District (Denver), State of Colorado; Member, John D. and Catherine T. MacArthur Foundation’s Research Network on Law and Neuroscience; Research Fellow, Gruter Institute for Law and Behavioral Research. The views expressed in this essay do not necessarily reflect the views of the Gruter Institute, the MacArthur Foundation or any of my colleagues on the Denver District Court.
Kudos to Gail and Richard Murrow for considering the potential links between pain, lack of empathy, dehumanization, and antisocial behavior.\(^1\) They are no doubt onto something deep and important about how the mysteries of pain and empathy—mysteries that seem at once central to our humanness yet puzzlingly inaccessible—might be better understood by trying to discover how those phenomena seem to be represented in the brain. Had they stopped there, this would be an important and engaging paper. But in their enthusiasm for finding some ‘policy’ relevance for these links, they exhibit the worst kind of what my friend Stephen Morse has called ‘brain overclaim syndrome’.\(^2\) Neuroscience cannot answer, and likely will never be able to answer, these kinds of normative policy questions.

I have a few smaller bones to pick with the paper before we get to the Brontosaurus of public policy. The paper is awash in what neuroscientists call the ‘reverse inference problem’.\(^3\) Identifying the neural networks associated, say, with empathy does not mean those networks ‘cause’ empathy as opposed to the other way round; indeed, the two might simply be correlated with no direct cause in either direction. This error can probably be forgiven, not only because the Murrows are quite up front in labeling these alleged links as ‘hypothetical’, but also because in recent years there have been some thoughtful proposals suggesting that there may indeed be some limited circumstances where the reverse inference might be appropriate.\(^4\)

The discussion of empathy, especially the assumption that empathy is a conspecific phenomenon—ie it is triggered only when we see pain or emotion in another human—seems quite oversimplified. Famous work by David Premack and his colleagues showed that human infants could recognize ‘hurting’ and ‘helping’ behaviors—and even showed a preference for ‘helping’ behaviors—even when the helping and hurting was being ‘done’ by geometric shapes.\(^5\) It seems that our brains come pre-equipped with a broadly generalized, and not at all conspecific, capacity for empathy. Anyone who melts at those abused pet commercials certainly knows that human empathy is not confined to other humans.\(^6\)

The paper’s simultaneous overreliance and underreliance on mirror neurons as the supposed medium of empathy is also puzzling. The Murrows say that mirror neurons are ‘believed’ to be present in humans, and for this proposition they cite a 2004 paper

---

1. G.B Murrow & R. Murrow, A Hypothetical Neurological Association Between Dehumanization and Human Rights Abuses, J. LAW & BIOSCI. 1, 29, June 8, 2015 [hereinafter referred to as ‘Murrows’].
2. Stephen J. Morse, Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note, 3 OHIO ST. J. CRIM. L. 397 (2006).
3. Geoffrey K. Aguirre, Functional Imaging in Behavioral Neurology and Cognitive Neuropsychology, in BEHAVIORAL NEUROLOGY AND COGNITIVE NEUROPSYCHOLOGY 85, 96 (Todd E. Feinberg & Martha J. Farah, eds., 2003).
4. Russell A. Poldrack, Can Cognitive Processes Be Inferred from Neuroimaging Data?, 10 TRENDS COGN. SCI. 59, 63 (2006).
5. David Premack, The Infant’s Theory of Self-Propelled Objects, 36 COGNITION 1, 16 (1990); David Premack & Ann J. Premack, Intention as Psychological Cause, in CAUSAL COGNITION: A MULTIDISCIPLINARY DEBATE 185 (Dan Sperber, David Premack & Ann J. Premack, eds., 1995).
6. The Murrows’ stated view is that this is not ‘real’ empathy, it is simply human empathy transferred to non-humans. This seems to be a distinction without any explanatory difference.
Neuroscience cannot answer these questions. But Roy Mukamel and his colleagues demonstrated in 2010 that humans have mirror neurons.

On the other hand, the Murrows assume that mirror neurons are the key to understanding empathy, even though in recent years their explanatory significance for many phenomena, including empathy, autism and psychopathy, has come under increasing criticism. The neuro-philosopher Patricia Churchland suggests that the mirror neuron explanation of empathy grossly oversimplifies the philosophical complexities of action, agency, and intentionality, and complains that ‘[c]ommon convictions about how mirror neurons explain skills in mental attribution have taken off in a way that leaves the evidence in the dust…’.

One of the paper’s weakest scientific moments is its reliance on the literature of implicit bias as an example of how all of us are saddled with biases driven by impulses to dehumanize. The authors needn’t have resorted to such a controversial and unsettled literature. Humans have been biased, and have largely been aware of those biases, since we first emerged 200,000 years ago. Indeed, ‘bias’ is an unfairly loaded word. Bias is just one part of our emotional toolbox, like jealousy and guilt, that served us well 200,000 years ago, and continues to serve us and disserve us, depending on the environmental context.

In fact, I was surprised that the paper glosses over the fact that the kinds of ‘bad’ dehumanizing biases that seem so problematical today—ethnic and racial biases—have their own deep evolutionary roots. The Murrows pay passing reference to the fact that ‘outgroups are universally distinguished as “not human” or as “less than human”’. But this story is much more profound than this offhanded comment might suggest.

Because we evolved in small groups of mostly related individuals, we have brains that powerfully discriminate between members of our own group and non-members. We have strong, albeit conflicting, prosocial feelings toward members of our group, but

7 Murrows, supra note 1, at 10; Luigi Cattaneo & Giacomo Rizzolati, The Mirror Neuron System, 27 ANN. REV. NEUROSCI. 169, 192 (2004).
8 Roy Mukamel, et al., Single Neuron Responses in Humans During Execution and Observations of Actions, 18 Curr. Biol. 431, 434 (2010).
9 Gregory Hickok, Eight Problems With the Mirror Neuron Theory of Action Understanding in Monkeys and Humans, 7 J. Cogn. Neurosci. 1229, 1243 (2009).
10 PATRICIA S. CHURCHLAND, BRAINTRUST 127 (2011).
11 The implicit bias literature is quite controversial. Criticisms of it include the claim that the standard implicit association test (‘IAT’) is not really detecting bias, that it has proved unreliable when the same subjects are tested at different times, and that even if it were reliably measuring hidden bias there is no established empirical link between the allegedly biased attitudes uncovered by the IAT and biased behaviors. See eg Hart Blanton & James Jaccard, Arbitrary Metrics in Psychology, 61 AM. PSYCHOL. 27, 41 (2006); Hal R. Arkes & Philip E. Tetlock, Attributions of Implicit Prejudice, or Would Jesse Jackson ‘fail’ the Implicit Association Test?, 15 PSYCHOL. INQUIRY 257, 278 (2004); Amy Wax & Philip Tetlock, We are All Racists at Heart, WALL STREET JOURNAL (June 9, 2011).
12 There are some biases about which we seem famously unaware, and which psychologists delight in studying. One such kind of unconscious bias is called ‘cognitive dissonance’, and is grounded on the model that when the brain is faced with two conflicting pieces of information it sometimes distorts one piece to reconcile the conflict. Jack W. Brehm, Postdecision Changes in the Desirability of Alternatives, 52 J. ABNORM. & SOC. PSYCHOL. 384, 389 (1956). More modernly, Amos Tversky and Dan Kahneman have shown that humans regularly and quite predictably miscalculate many kinds of risks. For example, Amos Tversky & Dan Kahneman, Judgment Under Uncertainty: Heuristics and Biases, 185 SCIENCE 1124, 1131 (1974).
13 Murrows, supra note 1, at 4–5.
Neuroscience cannot answer these questions

powerful and largely unrestrained feelings of distrust toward outsiders. Sometimes, that deep distrust of outsiders can still have modern utility.

Is a man a racist if, late at night in a remote part of a city, he walks to the other side of the street to avoid a group of young black men? Is the answer to this question the same if the avoider is aware that young black men in that community commit violent crimes at a rate significantly higher than young white men? Is the answer different if the avoider himself is black? Is any of this even 'bias'?

The Murrows’ penultimate hypothesis—that dehumanization is caused by reduced empathy—is largely tautological in the unexamined manner they have presented it. What do they mean by ‘dehumanization’ other than behaviors that in fact reflect a lack of empathy? The connection as they try to describe it is both under- and overinclusive. For example, many and perhaps most psychopaths lack normal levels of empathy—that’s one of the clinical criteria of psychopathy. But not every psychopath acts out his lack of empathy in acts of violence, or ‘human rights abuses’ as the Murrows put it. In fact, the majority of psychopaths are non-violent. Conversely, not every criminal, or even every human rights abuser, is a psychopath with reduced empathy. Human behavior is untidy, and these oversimplified categories do an injustice to our complexity.

And now to the Brontosaurus. The Murrows argue that the chain connecting dehumanization with human rights abuses, the links of which consist of neurological insights about pain and empathy, will guide us in negotiating some very difficult public policy issues. They take on two hard First Amendment issues: whether corporations should have free speech rights and whether hate speech crimes are unconstitutional. They contend that their hypothesized links argue in favor of answering both of these questions ‘no’. Here’s where the Murrows go seriously off the rails. With all due respect, I have a sneaking suspicion that neuroscience’s claimed answers to these difficult policy questions are actually the views the Murrows already had before they engaged in this exercise.

The Murrows label the association of African Americans with crime a ‘stereotype’, Id. at 6, but the sad fact is that the stereotype is true, at least when it comes to violent crime, and especially the two kinds of violent crime pertinent to our crossing-the-street hypothetical: homicide and robbery. In 2010, the year of the last census, African Americans comprised 13% of the US population. 2010 Census Quick Facts, http://www.census.gov/quickfacts/#table/PST045214/00 (accessed Oct. 7, 2015). But the FBI crime statistics for 2010 show that they were arrested for 39% of all violent crimes (a 300% overrepresentation), 52% of all homicides (400% overrepresentation) and 56% of all robberies (430% overrepresentation). US Department of Justice, Federal Bureau of Investigation, Crime in the United States 2010, Tbl. 43a, https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/tables/table-43 (accessed Oct. 7, 2015). Granted, these are arrest rates and not conviction rates, but national victim surveys have shown the arrest rates mirror the crime rates, and that since the mid-1970s US blacks have been disproportionately committing violent crime, and been the disproportionate victims of it. MICHAEL TONRY, MALIGN NEGLECT: RACE, CRIME AND PUNISHMENT IN AMERICA 4 (1995); Bureau of Justice Statistics, National Crime Victimization Survey 2012–2013, Special Tabulation.

In the Dec. 13, 1993 edition of Newsweek, Jesse Jackson famously said this: ‘There is nothing more painful to me at this stage in my life than to walk down the street and hear footsteps and start thinking about robbery, then look around and see somebody white and be relieved’.

ROBERT D. HARE, MANUAL FOR THE PSYCHOPATHY CHECKLIST–REVISED (2nd ed. 2003).

Scott Lilienfeld & Hal Arkowitz, What ‘Psychopath’ Means, in FORENSIC AND LEGAL PSYCHOLOGY 107 (Mark Constanzo & Daniel Krauss, eds., 2012).
Don’t get me wrong, I don’t necessarily think that it is a mistake to look to neuroscience for confirmation that our common wisdom about some issues is probably correct. If that’s what it takes to get people to recognize that humans come preequipped with certain evolved psychological and behavioral predispositions, then so be it. But it is one thing to say that we need neuroscience to help us be comforted in our folk wisdom (or even to talk us out of it, as the psychologies of cognitive dissonance and risk assessment have done in some narrow areas), and quite another to suggest that neuroscience will have any value in helping us answer, let alone be dispositive of, controversial policy questions. The Murrows could not have chosen two public policy issues more immune to a hypothetical neurological analysis than the two free speech issues they discuss.

First, these ‘policy’ questions are not really policy questions at all; they are legal questions about the meaning of our written constitution (which of course themselves may be part policy—more about that later). The Murrows may believe that the best policy is to prohibit corporations (and unions) from engaging in political speech and to criminalize certain kinds of hate speech, and indeed perfectly reasonable and thoughtful people support those policies. The Murrows may even think that neuroscience supports these conclusions. But no amount of neuroscience can change the words of the Constitution, unless it ends up being so persuasive that we amend those words (or, maybe just as unlikely, that it causes members of the Court to alter their interpretative approaches).

Of course, jurisprudence at the level of the United States Supreme Court is seldom just ‘calling balls and strikes’, as Chief Justice Roberts put it during his confirmation hearing. We all recognize that the issues that end up at the Supreme Court, especially the constitutional issues, often require difficult judgments not only about theories of interpretation and core constitutional principles, but also about how those core principles should apply to modern issues never dreamt about by the framers. But even the most extreme ‘living constitutionalist’ must admit that we do not write on an entirely blank slate, and are not free to go where the Murrows or their neuroscience proposes to take us if the voyage will violate the Constitution.

_Citizens United v. Federal Election Comm’n_19 was not really a First Amendment case, at least in the sense that most First Amendment cases require the Court to draw sometimes difficult substantive lines between protected behavior and unprotected behavior, between reasonable and unreasonable restrictions on free speech. Instead, the case depended on the threshold question of whether the protections of the First Amendment even apply to corporations. If they do, then the absolute corporate campaign prohibitions contained in the subject federal campaign finance statute20 were unquestionably unconstitutional. If the First Amendment does not apply, then the statute was unquestionably not unconstitutional, at least on free speech grounds. In answering this threshold question, the Court was not writing on a blank slate.

---

18 One of the best recent examples of this temporal disconnect is _Riley v. California_, 134 S.Ct. 2473 (2014), in which the Court held that the Fourth Amendment generally requires a warrant before police may search a cell phone.

19 558 U.S. 310 (2010).

20 Section 203 of the Bipartisan Campaign reform Act of 2002, 2 U.S.C. § 441b, prohibited corporations and unions from funding certain defined ‘electioneering communication’.
As of 1789, when the First Amendment was adopted, formal and informal associations of individuals had long engaged in political speech, and in mixed for profit and political behavior. This tradition included newspapers, which are expressly granted the ‘freedom of the press’ in the text of the First Amendment, and which also commonly did business (and, gasp, even made profits) in organized corporate or other associational forms, before and after the founding. The majority in *Citizens United* simply held that because there is nothing in the text of the First Amendment (which, by the way, does not actually use the word ‘person’ at all) to suggest the framers meant to exclude all of these long engaged-in activities from its purview, they are not excluded.\(^{21}\)

As I understand the Murrows’ argument about *Citizens United*, they posit that humans get into trouble not just when we fail to recognize the humanity of other humans, but also when we extend that humanity to non-humans, like corporations. *Citizens United* was wrongly decided. But by that grossly imprecise metric the Murrows should not object, at least on neuroscience-informed free speech grounds, if federal, state and local governments completely ban formally organized groups of like-minded protestors, say, the Socialist Workers Party 50 years ago or the Tea Party or Black Lives Matter today, from engaging in political speech.

Their argument that criminalizing hate crimes should have limited if any First Amendment implications is just as specious, both constitutionally and logically. On this point, I understand the Murrows’ neuroscience-informed argument to be this: because dehumanization is bad, any effort by the law to reduce dehumanizing behavior is good, and therefore should be constitutional. The core constitutional principle the Murrows’ syllogism fails to acknowledge is that the Constitution itself, and particularly the Bill of Rights that begins with the First Amendment, is all about restraining government, even when, and maybe especially when, it believes it is acting for the common good.

The difficult constitutional issue when it comes to hate crimes is not whether hate speech standing alone can be criminalized (it cannot),\(^{22}\) but rather whether the state may constitutionally increase the penalties for certain recognized crimes when hate is an additional element of a given crime. That is, I am perfectly free to call Al Sharpton the N-word, or Woody Allen the J-word, but I am not free to punch them while doing so. May the state constitutionally impose additional penalties on top of the penalties for assault, when the assault is accompanied by dehumanizing hateful words (or, even more controversially, unexpressed but equally hateful thoughts)?

Courts across the country have answered this question in many different ways. The general rule that has become largely settled is that the state may increase the penalties for a crime when it is motivated by animus evidenced by otherwise protected words, analogous to how it may increase the penalties for other criminal acts combined with otherwise protected actions, like aggravated robbery by a defendant using a gun he is licensed and authorized to carry. But the state may not criminalize protected speech standing alone. Neuroscientific insights into how hurtful some speech might be to some listeners will never reduce, and indeed arguably should increase, the protective vigilance required by the First Amendment.

---

\(^{21}\) There was a much more difficult issue in *Citizens United*: whether the Court should continue to follow what the majority characterized as an outlier case from 1990, *Austin v. Michigan Chamber of Commerce*, 494 U.S. 652, which for the first time held that political speech may be banned based on the speaker’s corporate identity.

\(^{22}\) *R.A.V. v City of St. Paul*, 505 U.S. 377 (1992).
Other hate-crime issues remain unsettled, including whether there is a point at which the mere ‘increase’ in penalty might be so severe that it should be considered a separate, unconstitutional, criminalization of the speech, and the extent to which improper motivation alone, perhaps evidenced by prior bad acts, is sufficient to prove a hate-crime aggravator. Many of these free speech cases also turn on how seriously courts consider the ‘chilling effect’ that regulation aimed at once set of unprotected behavior might have on other, protected, behavior. These are all thorny issues on which reasonable people can disagree. But I am afraid the axiom that dehumanization is bad will not answer any of these difficult questions. After all, spending time in jail for hurtful speech is no doubt also quite dehumanizing.

There are incredibly important questions to be asked about the relationships between human pain, empathy, violence, blame, punishment, and forgiveness, and the Murrows ask some of them. Future neuroscience may even inform some of these mysteries to the point it has some impact on policy. But current neuroscience has virtually nothing to say about the Constitution or public policy, and in hypothesizing otherwise the Murrows have done a disservice to the integrity of science, to the ideal of neutral law in a constitutional republic, and to the complexity of the human condition.