

The Science of Evil: What Do *You* Think?

Social Psychology

Name _____

Date _____

Period _____

In this assignment, you will be comparing several points of view when it comes to explaining why people do “evil” things. They are:

<p>A. Philip Zimbardo, the psychologist who ran the Stanford Prison Experiment</p> <ul style="list-style-type: none">○ Newspaper opinion piece “Power Turns Good Soldiers into ‘Bad Apples’”	<p>B. Simon Baron-Cohen, a psychologist who writes about the subject in his book, <i>The Science of Evil</i></p> <ul style="list-style-type: none">○ Essay “The Science of Evil” for The Huffington Post http://huff.to/VOCFRS
<p>C. Contemporary neuroscientists</p> <ul style="list-style-type: none">○ Summarized by journalist Ron Rosenbaum in the article “The End of Evil” on Slate.com	

You will organize these points of view in the boxes below. In each box, 1) in the space provided, make some **notes** about that point of view, then 2) summarize it in **one sentence**.

Philip Zimbardo

One sentence summary:

Simon Baron-Cohen

One sentence summary:

Power Turns Good Soldiers into 'Bad Apples'

Social Psychology

The Boston Globe

May 9, 2004

By Philip G. Zimbardo

The horrifying photos of young Iraqis abused by American soldiers have shocked the world with their depictions of human degradation, forcing us to acknowledge that some of our beloved soldiers have committed barbarous acts of cruelty and sadism. Now there is a rush to analyze human behavior, blaming flawed or pathological individuals for evil and ignoring other important factors. Unless we learn the dynamics of "why," we will never be able to counteract the powerful forces that can transform ordinary people into evil perpetrators. Those responsible should suffer severe sanctions if found guilty. However, we must separate guilt from blame. Should these few Army reservists be blamed as the "bad apples" in a good barrel of American soldiers, as our leaders have characterized them? Or are they the once-good apples soured and corrupted by an evil barrel? I argue for the latter perspective after having studied the psychology of evil for many decades. In fact, I have been responsible for constructing evil barrels that produced many bad apples.

Like Brigadier General Janis Karpinski, who was in charge of the Iraqi prison at Abu Ghraib, I was once a prison superintendent with no experience or training in corrections. In 1971 I was in charge of the Stanford Prison Experiment, in which randomly assigned student volunteers in a simulated prison role-played prisoners and guards. Although everyone knew it was just an experiment, the line between simulation and reality was breached as it became a psychological prison of incredible intensity.

The planned two-week study was terminated after only six days because it was out of control. Good boys chosen for their normalcy were having emotional breakdowns as powerless prisoners. Other young men chosen for their mental health and positive values eased into the character of sadistic guards inflicting suffering on their fellow students without moral compunction. And those "good guards" who did not personally debase the prisoners failed to confront the worst of their comrades, allowing evil to ripen without challenge.

The terrible things my guards did to their prisoners were comparable to the horrors inflicted on the Iraqi detainees. My guards repeatedly stripped their prisoners naked, hooded them, chained them, denied them food or bedding privileges, put them into solitary

confinement, and made them clean toilet bowls with their bare hands.

As the boredom of their job increased, they began using the prisoners as their playthings, devising ever more humiliating and degrading games for them to play. Over time, these amusements took a sexual turn, such as having the prisoners simulate sodomy on each other. Once aware of such deviant behavior, I closed down the Stanford prison.

Human behavior is much more under the control of situational forces than most of us recognize or want to acknowledge. In a situation that implicitly gives permission for suspending moral values, many of us can be morphed into creatures alien to our usual natures. My research and that of my colleagues has catalogued the conditions for stirring the crucible of human nature in negative directions. Some of the necessary ingredients are: diffusion of responsibility, anonymity, dehumanization, peers who model harmful behavior, bystanders who do not intervene, and a setting of power differentials.

Those factors were apparently also operating in Iraq. But in addition there was secrecy, no accountability, no visible chain of command, conflicting demands on the guards from the CIA and civilian interrogators, no rules enforced for prohibited acts, encouragement for breaking the will of the detainees, and no challenges by many bystanders who observed the evil but did not blow the whistle.

We must learn from this tragic event so it is never repeated. And we must not permit the authorities to deflect the blame and responsibility from themselves by pointing fingers at those soldiers who went into the administration's preemptive war as proud Americans and return now as disgraced prison guards.

The arrogance of power that spawned the "shock and awe" of military might one short year ago has been humbled by the dismay and disgust over these revelations of abuse. It is time for all Americans to reflect on the justification for continuing the war in Iraq that is killing, maiming, and demeaning our young men and women who have been put in harm's way for spurious reasons.

Before more of our youth are corrupted, perhaps the time has come to empty out the vinegar of needless war that has filled that evil barrel.

The Science of Evil

Social Psychology

http://www.huffingtonpost.com/simon-baroncohen/science-of-evil_b_2831311.html

By Simon Baron-Cohen
3/8/13

Phil Zimbardo's TEDTalk on evil is a terrific example of how to give a good TEDTalk, and a terrific example of the situationist explanation for how ordinary people do 'evil' things under the wrong conditions. These conditions, according to Zimbardo, include anonymity (e.g., when soldiers wear uniforms), obedience to an authority figure who says they will take responsibility, being given power to use and abuse, and the absence of any moral spotlight to supervise what people are up to. Zimbardo is no doubt right that such situational factors can turn "good apples" into "bad apples" and -- in his metaphor -- how one needs to look at the "bad barrel" to explain how people cross the line from being a good person to a bad person.

Zimbardo uses the term 'evil' but I prefer a different term, the 'erosion of empathy', because whilst the word 'evil' is used as if it is an explanation ("He did X because he is evil"), in fact evil is just a word meaning "the absence of good", getting us into a dangerous circularity ("He did X because he is not good"). The term 'evil' doesn't take us any further forward. In contrast, the concept of empathy (or its erosion, by degrees) has explanatory power, because empathy can be measured and analysed scientifically in terms of the factors that influence how much empathy each of us has. Even more usefully, empathy is normative, meaning we don't just have to focus on the extreme negative end of human behavior. We can consider the whole spectrum, from cruelty through to kindness, and every shade in between.

I enjoyed Zimbardo's TEDTalk not just because it was packed full of interesting photography and footage, but also because of the personal anecdotes: How Zimbardo was in the same high school class as Stanley Milgram, who ran the other world-famous social psychology experiment to explain cruelty. And how the young female whistle-blower in the Stanford Prison Experiment, who persuaded Zimbardo to stop the experiment because it was damaging the young volunteer-prisoners, later became his wife!

Equally uplifting was how Zimbardo now uses his same situationist approach to explain what he neatly calls "the banality of heroism," adapting Hannah Arendt's chilling phrase "the banality of evil," to argue that ordinary people can choose the path of helping someone or can

simply be passive and let bad things happen. He picks out Wesley Autrey as an example of an ordinary hero. This was the black construction worker who saw a young white student having an epileptic fit, causing him to fall onto the train tracks as a New York subway train was approaching. Wesley gave his two young daughters to a stranger to look after, and made a split-second decision to dive onto the train track, cover the epileptic student with his body, holding him against the ground, allowing the train carriages to pass over them both, and missing his head by just half an inch.

Missing from Zimbardo's account is any reference to the individual factors that interact with such situational ones that make one person conform to socially sanctioned group violence, and make another person resist the conformity bias and stand up for what he or she thinks is right. - Simon Baron-Cohen

But whilst I applaud Zimbardo for keeping alive the situationist theory of cruelty for over four decades, I think the theory is a product of its time. When the Stanford Prison Experiment was conducted, in 1971, it was fashionable to 'blame' the ills of the world on 'the system' (our social institutions).

Missing from Zimbardo's account is any reference to the individual factors that interact with such situational ones that make one person conform to socially sanctioned group violence, and make another person -- in the very same situation -- resist the conformity bias and stand up for what he or she thinks is right. Zimbardo hints at individual factors when he says we each have a "choice" to do the right thing or not, but that does not take us very deep and again is even a little circular. We cannot explain cruelty or kindness simply as the result of making choices, since what we are trying to explain is why people make the choices they do.

Research over these last four decades has isolated a number of factors that bias individuals in one direction or another. The first is the quality of attachment in early experience. Research into adults with personality disorders reveal high rates of abuse and emotional neglect in their early childhoods, suggesting that it is the experience of being cared for as an infant and toddler that influences one's capacity to care about others in later life.

Early experience can't be the whole story, since not everyone who has a bad childhood ends up showing

The Science of Evil

Social Psychology

http://www.huffingtonpost.com/simon-baroncohen/science-of-evil_b_2831311.html

cruelty to others, and some psychopaths had perfectly loving parents and still ended up capable of hurting others. This means we have to turn to biology. Research again points to our genes influencing our risk for developing antisocial personality traits. For example, Avshalom Caspi and colleagues found that which version of the MAO-A gene you carry affects your risk of ending up with a diagnosis of "conduct disorder" or delinquency in adolescence. Most interesting, this genetic variant interacts with whether a person experienced abuse or neglect in their childhood, demonstrating that genes interact with environment.

And a second biological factor that influences later empathy is the prenatal sex steroid hormone testosterone. How much of this hormone the fetus is exposed to shapes the development and function of the 'empathy' circuit in each of our brains. My own TED talk acknowledges the major contribution of Zimbardo's and Milgram's social psychological theories of cruelty, but reminds us that we need to understand its neurobiological and developmental psychological roots too.

Simon Baron-Cohen is author of The Science of Evil (Basic Books) and is Professor and Director of the Autism Research Centre, Cambridge University.

The End of Evil?

Social Psychology

By Ron Rosenbaum

Sept. 30, 2011

Neuroscientists suggest there is no such thing. Are they right?

Is evil over? Has science finally driven a stake through its dark heart? Or at least emptied the word of useful meaning, reduced the notion of a numinous nonmaterial malevolent force to a glitch in a tangled cluster of neurons, the brain?

Yes, according to many neuroscientists, who are emerging as the new high priests of the secrets of the psyche, explainers of human behavior in general. A phenomenon attested to by a recent torrent of pop-sci brain books with titles like *Incognito: The Secret Lives of the Brain*. Not secret in most of these works is the disdain for metaphysical evil, which is regarded as an antiquated concept that's done more harm than good. They argue that the time has come to replace such metaphysical terms with physical explanations—malfunctions or malformations in the brain.

Of course, people still commit innumerable bad actions, but the idea that people make conscious decisions to hurt or harm is no longer sustainable, say the new brain scientists. For one thing, there is no such thing as "free will" with which to decide to commit evil. (Like evil, free will is an antiquated concept for most.) Autonomous, conscious decision-making itself may well be an illusion. And thus intentional evil is impossible.

Have the new neuroscientists brandishing their fMRIs, the ghostly illuminated etchings of the interior structures of the skull, succeeded where their forebears from disciplines ranging from phrenology to psychoanalysis have failed? Have they pinpointed the hidden anomalies in the amygdala, the dysfunctions in the prefrontal lobes, the electrochemical source of impulses that lead a Jared Loughner, or an Anders Breivik, to commit their murderous acts?

And in reducing evil to a purely neurological glitch or malformation in the wiring of the physical brain, in eliminating the element of freely willed conscious choice, have neuroscientists eliminated as well "moral agency," personal responsibility? Does this "neuromitigation" excuse—"my brain made me do it," as critics of the tendency have called it—mean that no human being really wants to do ill to another? That we are all innocent, Rousseauian beings, some afflicted with

defects—"brain bugs" as one new pop-neuroscience book calls them—that cause the behavior formerly known as evil?

Are those who commit acts of cruelty, murder, and torture just victims themselves—of a faulty part in the head that might fall under factory warranty if the brain were a car?

The new neuroscience represents the latest chapter in a millennia-old and still divisive cultural conflict over the problem of evil, the latest chapter in the attempt by science to reduce evil to malfunction or dysfunction rather than malevolence. It's a quest I examined in *Explaining Hitler: the way the varieties of 20th-century psychological "science" sought to find some physiological, developmental, sexual, or psychoanalytic cause for Hitler's crimes*. (One peer-reviewed paper sought to trace Hitler's evil to a mosquito bite—to the secondary sequelae of mosquito-borne encephalitis which were known to cause profound personality changes as long as a decade after being contracted in the trenches of World War I.)

It would be consolatory if not comforting if we could prove that what made Hitler Hitler was a malfunction in human nature, a glitch in the circuitry, because it would allow us to exempt "normal" human nature (ours for instance) from having Hitler potential. This somewhat Pollyannaish quest to explain the man's crimes remains counterintuitive to many. I recall the late British historian and biographer of Hitler Alan Bullock reacting to the claims of scientism by exclaiming to me vociferously: "If he isn't evil, then who is? ... If he isn't evil the word has no meaning."

Indeed recent developments demonstrate that evil remains a stubborn concept in our culture, resistant to attempts to reduce it to pure "physicalism." To read the mainstream media commentary on the Breivik case, for instance, is to come upon, time after time, the word "evil." Not just that the acts were evil, but that he, Breivik was, as a *Wall Street Journal* columnist put it, "evil incarnate."

But what exactly does that mean? The incarnation of what? Satan? The word "incarnation," even without explicit religious context, implies, metaphorically at least, the embedding of a metaphysical force in a physical body. One can understand the scientific aversion to this as a description of reality. But evil as a numinous force abides. It is not surprising that Pope Benedict issued a

The End of Evil?

Social Psychology

statement following the attacks in Norway calling on everyone to "escape from the logic of evil." (Although what exactly is that "logic"?)

Even if it was not surprising for the Pope to invoke evil thus, it was surprising to see a devout atheist such as my colleague Christopher Hitchens invoke "evil" in his "obituary" for Osama bin Laden. Hitchens admits wishing he could avoid using "that simplistic (but somehow indispensable) word." But he feels compelled to call whatever motivated bin Laden a "force" that "absolutely deserves to be called evil."

But what is this "force," which sounds suspiciously supernatural for an atheist to believe in? Some kind of Luciferian Kryptonite? Where is it located: in the material or nonmaterial world?

That is the real "problem of evil" (or, to use the technical term philosophers employ for conscious, freely-willed, evil-doing: "wickedness"). We tend to believe it exists: Popular culture has no problem with it, giving us iterations from Richard III to Darth Vader; politicians use it promiscuously ("the axis of evil"). But even religious thinkers continue to debate what it is—and why a just and loving God permits evil and the hideous suffering it entails to prevail so often, or even—if they shift the blame to us (because God gave man free will to sin)—why God couldn't have created a human nature that would not so readily choose genocide and torture. (For the record, I'm an agnostic.)

This argument has been going on for more than a millennium, at least since Augustine proclaimed that evil was in the realm of "non-being," which seems to some a great evasion. Meanwhile pop neuroscience—and its not-very-well-examined assumptions—has taken center stage in the struggle to put evil in its place under the thumb of science.

One person whose work on these matters has received considerable attention lately is the British Professor of Psychopathology, Simon Baron-Cohen. (Yes, cousin of Sacha Baron-Cohen aka Borat, but highly regarded as a serious scientist.) He's the author of *The Science of Evil*, which seeks to dispose of the problem of evil in part at least by changing its name.

"My main goal," says Baron-Cohen, "is replacing the unscientific term 'evil' with the scientific term 'empathy.' "What he means is that instead of calling someone evil we should say they have no empathy.

Baron-Cohen goes to great lengths to posit an "empathy circuit" in the brain whose varying "degrees" of strength constitute a spectrum, ranging from total, 100 percent empathy to "zero degrees of empathy."

This empathy circuit, he tells us, consists of 13 specific regions of the brain involved in the generation of nonevil choices, among them "the medial prefrontal cortex," "the inferior frontal gyrus," and "the posterior superior temporal sulcus."

Ideally all of these act together empathetically to defeat "single minded focus," which appears to be Baron-Cohen's explanation for what was previously called evil. Single-mindedness is the inability to "recognize and respond" to the feelings of others. A healthy empathy circuit allows us to feel others' pain and transcend single-minded focus on our own. This theory does, however, seem to carry a presumption that when one "recognizes and responds," one will do so in warm and fuzzy ways. But what about those who "recognize and respond" to others' feelings with great discernment—and then torture them? It happens.

One troubling aspect of Baron-Cohen's grand substitution of a lack of empathy for evil is the mechanistic way he describes it.

He characterizes those who lack empathy as having "a chip in their neural computer missing." He tells us "empathy is more like a dimmer switch than an all-or-none switch." The big problem here is that by reducing evil to a mechanical malfunction in the empathy circuit, Baron-Cohen also reduces, or even abolishes, good. No one in this deterministic conceptual system chooses to be good, courageous, or heroic. They just have a well-developed empathy circuit that compels them act empathetically—there's no choice or honor in the matter.

And so evil for Baron-Cohen is just "zero degrees of empathy." And I'm left with the nonempathetic feeling that his boast that he is "replacing" evil with nonempathy is more a semantic trick than a scientific discovery. It's another instance of what one of the authors in an important collection of academic papers from MIT Press called *Neuroethics*, calls "Brain Overclaim Syndrome."

A number of papers in *Neuroethics* pour cold water on the triumphalism of the giddy new pop-sci brain books.

The End of Evil?

Social Psychology

It makes clear there is a debate within the neuroscience profession about what exactly all those impressive-looking fMRI images tell us. And these "neurocritics" or "neuroskeptics" warn about the consequences for acting too quickly on these claims. (There is a valuable British website called Neuroskeptical that offers the general reading public these critiques and correctives from the point of view of someone within the profession. People need to know!)

The "Brain Overclaim" paper by Stephen Morse of the University of Pennsylvania's Center for Neuroscience and Society is a tongue-in-cheek "diagnostic note" on the grandiosity of the assumptions of the brain-book fad, mainly concerned about the way they have been creeping into jurisprudence. fMRIs have made their way into a Supreme Court opinion this year, for instance; Justice Stephen Breyer cited "cutting edge neuroscience" in his dissent to a ruling denying the right of California to ban violent video games, because the otherwise-pro-free-speech justice was alarmed at neuroscientific studies that claim such games could create mental pathways for actual violence.

But Morse's critique extends beyond the jurisprudential and goes to the heart of the failure of current neuroscience to explain or "replace" evil. Popular neuroscience has claimed to find the neural locus of love and God and evil, but Morse points out a fundamental flaw in their logic:

Despite all the astonishing advances in neuroscience, however, we still know woefully little about how the brain enables the mind and especially about how consciousness and intentionality can arise from the complicated hunk of matter that is the brain. ... Discovering the neural correlates of mental phenomena does not tell us how these phenomena are possible.

In other words, correlation doesn't always equal causation: We may know the 13 regions that light up on an fMRI when we feel "empathy" (or fail to light up when we choose evil) but that doesn't explain whether this lit-up state indicates they are causing empathy or just reflecting it.

The problem of evil—and moral responsibility—is thus inseparable from what is known in the philosophical trade as "the hard problem of consciousness." How does the brain, that electrified piece of meat, create the mind and the music of Mozart, the prose of Nabokov? Where is consciousness, anyway?

Many neuroscientists, confronted by the "hard problem of consciousness," evade it by citing a quarter-century-old experiment by one Benjamin Libet, which purported to reveal that apparently conscious decisions are actually made unconsciously—preconsciously—some 500 milliseconds (half a second) before the illusion of a conscious decision is made conscious. (A recent paper puts it at a full second.) But Libet's study fails to explain how the initial unconscious decision is made by the electrified piece of meat—he just kicks the can into the preconscious, you might say—or why we have the illusion of consciousness at all. It does suggest that those who purport to study the science of the brain do themselves—and science—a disservice by failing to learn from the contexts of history, logic, and very basic philosophy.

Those neuroscientists who disdain the idea of consciousness or free will and believe that Libet has disproved it all ought at least to give some attention to Francis Crick. Crick, whose co-discovery of DNA earned him a Nobel Prize and who recently daringly proposed a scientific locus for free will, offers his candidate for its neural coordinates. In his 1994 study "The Astonishing Hypothesis" Crick places it somewhere in or near the area called "the anterior cingulate sulcus" which is "next to Brodman's area 24. This is on the inside surface [of the skull] ... toward the front ... and near the top" of the brain. If that's the center of free will it's the center of evil as well. But even if Crick has trumped Libet, neither has dealt with the most disturbing implications of the new research that purports to find neural explanations for evil.

One can find some of these troubling possibilities laid out in a paper by Jonathan Marks of Harvard's Safrin Center for Ethics and Pennsylvania State University in the *American Journal of Bioethics*. The paper is called "A Neuroskeptical Guide to Neuroethics and National Security," and in it Marks references a growing resistance to "brain over-claims" within the profession. His objections are technical and ethical. He criticizes both the fetishizing of fMRIs, and their misuse. He reminds laymen looking at all the impressive fMRIs in pop-psych brain books that they are not actual images of individual brains in action, but rather composites based on statistical compilations of images of multiple brains, overlaid with special effects lighting he compares to "Doppler-weather radar images."

The End of Evil?

Social Psychology

"Would it be going too far to call this Photoshopping?" I asked Marks in a phone conversation.

"Photoshopping isn't the right word, but in one sense, it doesn't go far enough," he said. The images are "constructed from the start."

Marks' paper warns of "aggressive marketing" of fMRI scans by intelligence-contractor types as "lie detector" substitutes that could be used to select candidates for "enhanced interrogation" if their fMRI indicates potential deception under ordinary interrogation.

And he offered what I thought was one of the wisest responses to the debate over the existence of evil (and thus free will):

What he suggested is that we ought to act as if we had free will to choose good or evil.

And his warnings against the consequences of believing otherwise are validated by the fantasies of some fMRI enthusiasts. Consider, for instance, one of the more prominent new brain books: David Eagleman's *Incognito*.

In an excerpt in the Atlantic's "big ideas" issue, Eagleman depicts an Orwellian future in which fMRI scans will be used to preemptively identify those who have the potential to commit acts formerly known as evil, and prescribes for such possible malfeasants a regimen of "prefrontal workout[s]" to "better balance" those selected (how? by whom?) for brain remodeling.

He actually goes so far as to say, "Some people will need to be taken off the streets," on the basis of their fMRIs, "for a longer time (even a life time)." Neuroscientific totalitarianism invades your brain! The ultimate panopticon. No one seemed to notice or to care. It's science!

No mention of constitutional rights or preemptive detention or the Orwellian implications of this for radical dissenters, say, those whose rage against injustice might need to be toned down in the brain gyms.

I hesitate to say it, but these are evil ideas. Indeed, reading Eagleman, and returning to this debate about evil, led me to think about something that had occurred to me in examining the fallacious attempts to scientize Hitler. Evil does not necessarily inhere in some wiring diagram within the brain. Evil may inhere in bad ideas,

particularly when they're dressed up as scientific (as Hitler did with his "scientific racism").

As for evil itself, the new neuroscience is unlikely to end the debate, but it may cause us to be more attentive to the phenomenon. Perhaps evil will always be like the famous Supreme Court pronouncement on pornography. You know it when you see it. I don't like its imprecision, but I will concede I don't have a better answer. Just that we can do better than the mechanistic, deterministic, denial of personal responsibility the neuroscientists are offering to "replace" evil with.

I recall an exchange in my conversation with one of the original neuroskeptics, Daniel S. Reich, now head of a research division on nerve diseases at the National Institutes of Health. Reich was one of the first to critique "neuromarketing"—the promotion of fMRI technology to help pushers of commercial products and political candidates learn what words and images lit up what buttons in the brains of consumers and voters.

Toward the end of our conversation I asked Reich if he believed in evil. He was silent for a bit and then started talking about Norway. About degrees of evil. About the difference between the typical suicide bomber and the Oslo killer. How the former has only to press a button to accomplish his murderous goal and never has to see the consequences.

But on that summer camp island in Oslo, Reich said, Breivik was stalking victims for hours. He'd shoot one or more and, according to survivors, not register anything, just continue trudging forward, looking for more.

"He saw the consequences, the blood gushing, heard the screams. He just kept going." Some will try to say this is sociopathy or psychopathy or zero degrees of empathy and other exculpatory cop-outs. But fueled by his evil ideas Breivik kept going. To echo Bullock, if we can't call him evil who can we?