Speculations on the Evolutionary Origins of System Justification

John T. Jost¹, Robert M. Sapolsky², and H. Hannah Nam³

Abstract
For centuries, philosophers and social theorists have wondered why people submit voluntarily to tyrannical leaders and oppressive regimes. In this article, we speculate on the evolutionary origins of system justification, that is, the ways in which people are motivated (often nonconsciously) to defend and justify existing social, economic, and political systems. After briefly recounting the logic of system justification theory and some of the most pertinent empirical evidence, we consider parallels between the social behaviors of humans and other animals concerning the acceptance versus rejection of hierarchy and dominance. Next, we summarize research in human neuroscience suggesting that specific brain regions, such as the amygdala and the anterior cingulate cortex, may be linked to individual differences in ideological preferences concerning (in)equality and social stability as well as the successful navigation of complex, hierarchical social systems. Finally, we consider some of the implications of a system justification perspective for the study of evolutionary psychology, political behavior, and social change.

Keywords
system justification, ideology, political neuroscience, amygdala, hierarchy, evolutionary psychology

Date received: September 18, 2017; Accepted: February 18, 2018

We, as a species, have an easy time convincing ourselves and each other that our current ethic, morality, and way of living is not just good or better but comes close to being Truth. We, today, regard as barbarous such practices as selective infanticide, sacrifices to pagan gods, or the enslavement of human beings—practices that in earlier times we would have deemed just, honorable, and morally right. The things we do are always good.

Leon Festinger (1983, p. 163)

In the middle of the 16th century, a 22-year-old law student in France by the name of Etienne de la Boétie wrote an essay entitled Discourse of Voluntary Servitude that was to be circulated among academics for centuries to come (Lukes, 2011; Rothbard, 1975/2008; Stanley, 2015). In this work, de la Boétie (1548/2008) set out to understand:

how it happens that so many men, so many villages, so many cities, so many nations, sometimes suffer under a single tyrant who has no other power than the power they give him; who is able to harm them only to the extent to which they have the willingness to bear with him . . . . Surely a striking situation! Yet it is so common that one must grieve the more and wonder the less at the spectacle of a million men serving in wretchedness, their necks under the yoke, not constrained by a greater multitude than they, but simply, it would seem, delighted and charmed by the name of one man alone whose power they need not fear. (pp. 40–41)

The author defended three major hypotheses concerning the politics of obedience. According to Lukes (2011, p. 20), these amount to (a) “cultural inertia” or the “force of custom and habit”; (b) “manufactured consent,” that is, ideology and propaganda; and (c) “patronage,” such that “tyrants surround themselves with dependents, who in turn have their own dependents.” The anthropological record suggests that religious rituals, including human sacrifices, also “played a powerful role in the construction and maintenance of stratified

1 Corresponding Author: John T. Jost, Department of Psychology, New York University, 6 Washington Pl., New York, NY 10003, USA. Email: john.jost@nyu.edu

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
societies” (Watts, Sheehan, Atkinson, Bulbulia, & Gray, 2016, p. 3) in early civilization by “combining displays of ultimate authority—the taking of a life—with supernatural justifications that sanction authority as divinely ordained” (p. 10).

In the 450 years since de la Boétie’s *Discourse*, a celebrated cadre of intellectuals—including David Hume, Leo Tolstoy, Henry David Thoreau, Wilhelm Reich, Hannah Arendt, Michel Foucault, and Vaclav Havel—have revisited the fundamental questions he raised about why people submit willingly, even enthusiastically, to the humiliations of the powerful, and the conclusions they reached echoed those of de la Boétie (e.g., Lukes, 2011; Rothbard, 1975/2008; Stanley, 2015). Several commentators have noted that de la Boétie’s description of voluntary servitude (sometimes referred to as “self-domination”) has much in common with Marxist concepts of ideological hegemony and false consciousness (e.g., Rosen, 1996). To take just one example, the famous Italian Marxist, Gramsci (1971), marveled at the “spontaneous” consent given by the great masses of the population to the general direction imposed on social life by the dominant fundamental group” and proposed that “this consent is ‘historically’ caused by the prestige (and consequent confidence) which the dominant group enjoys because of its position and function in the world of production” (p. 12).

Political economists today marvel at the so-called paradox of inequality—the fact that despite sharp increases in economic disparities between the rich and poor over the past several decades, concerns about inequality seem to be on the decline, while citizens become more and more convinced that income differences are justified in terms of meritocratic considerations such as hard work, talent, and ambition (e.g., Kelly & Enns, 2010; Luttig, 2013; McCall, 2013; Mijs, 2017; Paskov & Dewilde, 2012). A plausible explanation is that people living in capitalist societies grow increasingly tolerant of inequality as way of justifying—and, indeed, coping with—harsh economic realities (e.g., Bénabou & Tirole, 2006; Jost, Banaji, & Nosek, 2004; Trump, in press). Even unambiguously oppressive social systems—such as slavery, caste systems, segregation, apartheid, and patriarchy—withstood shockingly long periods of stability and even perceived legitimacy before concerted efforts to overthrow them were finally undertaken.

Although it is true that people sometimes do turn against unjust manifestations of authority (Gurr, 1970), the persistence of inequality and exploitation leads social historians such as Zinn (1968/2002) to conclude that “Rebellion is only an occasional reaction to suffering in human history; we have infinitely more instances of forbearance to exploitation, and submission to authority, than we have examples of revolt” (p. 16). The question is why people are as accepting as they are of social injustice. Even if we allow that the acquiescence of some can be bought with “patronage” and cynical concessions to material self-interest (Bueno de Mesquita & Smith, 2011), human beings are, among many other things, “ideological animals” (Jost, Fitzsimons, & Kay, 2004), and ideology plays a meaningful role in the affect, cognition, motivation, and behavior of a great many individuals and groups (Jost, 2006). For a variety of reasons, then, people may internalize the norms of the social order on which they depend (Fehr & Gintis, 2007), and in so doing develop “mental resistance to the fundamental flaws of their social order” (Kuran, 1991, p. 32).

### A Theory of System Justification

It is hardly surprising that de la Boétie’s (1548/2008) student essay, penned during the Renaissance period, falls short of providing a complete or adequate theory of how and why human beings submit to tyrannical regimes (e.g., Gunn & Wilding, 2012; Rosen, 1996). Nevertheless, some of his observations about human nature—along with Gramsci’s (1971) emphasis on the popular tendency to experience “the existing social order” as a “stable, harmoniously coordinated system”—anticipate the framework of system justification theory, a social–psychological perspective that seeks to elucidate the individual and group-level mechanisms contributing to false consciousness and related ideological phenomena (Jost & Banaji, 1994; Jost & van der Toorn, 2012). According to system justification theory, people are motivated—often at a non-conscious level of awareness—to defend, bolster, and justify the social, economic, and political institutions on which they depend (Jost, Banaji, et al., 2004; Jost & Hunyady, 2005; Liviatan & Jost, 2014). Thus, as Veyne (1976/1992) put it, “the tendency to justify what exists constitutes one of the factors which combine to shape opinions” (p. 379, emphasis added), including opinions about the legitimacy of hierarchy.

Empirical studies in support of system justification theory demonstrate that when women, for instance, are made to feel dependent upon the social system—or are exposed to criticisms of the system—they come to view gender disparities in politics and business as natural, desirable, and just (Kay et al., 2009). In other cases, women may regard themselves as little more than sexual objects existing for the gratification of men (Calogero & Jost, 2011). Interviews with domestic workers in post-Apartheid South African homes reveal that—far from seeing themselves as underpaid or exploited—these women, most of whom were Black, saw themselves as lucky to be part of a symbiotic relationship with their wealthy White employers (Durrheim, Jacobs, & Dixon, 2014). And, rather than blaming their problems on the social system, low-income Latina and African American mothers in the United States attributed poverty to drug and alcohol addiction and “character deficiencies of the poor” (Godfrey & Wolf, 2016). Despite significant disparities in income, education, employment, and health, low-status minorities in New Zealand (Maori, Asians, and Pacific Islanders) legitimize ethnic group relations as much as, if not more than, members of the European majority (Sengupta, Osborne, & Sibley, 2015).

From a system justification perspective, endorsing politically conservative ideologies—that is, ideas that lend legitimacy to the traditional social, economic, and political order—yields palliative effects such as increased satisfaction with the status quo (Jost & Hunyady, 2002, 2005). This helps to explain why conservatives and rightists in the United States and
Europe report being happier than liberals and leftists, who are more troubled by putative injustices, such as high levels of social or economic inequality (Napier & Jost, 2008; see also Bixter, 2015; Burton, Plaks, & Peterson, 2015; Butz, Kieslich, & Bles, 2017; Choma, Busseri, & Sadava, 2009; Cichocka & Jost, 2014; Onraet, Van Asche, Roets, Haesevoets, & Van Hiel, 2016). This does not mean that conservatives are objectively happier or healthier than others (see Wojciszke, Hovasapian, Graham, Motyl, & Ditto, 2015)—or that conservative societies improve well-being (Okulicz-Kozaryn, Holmes, & Avery, 2014). On the contrary, rationalization and denial are at the very heart of system justification. These ideas are consistent with Flannery and Marcus’s (2012) observation that human “emotions . . . play a role in the subordination of our self-interest for the good of the group” (p. 561).

Because it is painful to acknowledge that one is living in a state of injustice or exploitation (Barclay & Saldanha, 2015; Janoff-Bulman, 1992; Lerner, 1980), those who are disadvantaged may be motivated to reach the conclusion that things are not really as bad as they seem (Laurin, Fitzsimons, & Kay, 2011). This process of rationalization may provide hedonic benefits for the individual, but it also undermines support for collective action (Jost, Wakslak, & Tyler, 2008; Sengupta, Greaves, Osborne, & Sibley, 2017; Wakslak, Jost, Tyler, & Chen, 2007). That is, individuals—including members of disadvantaged groups—who defend and bolster the legitimacy of the social system are less willing to protest on behalf of the disadvantaged, in comparison with those who question the system’s legitimacy (Becker & Wright, 2011; Jost, Becker, Osborne, & Badaan, 2017; Jost et al., 2012; Osborne & Sibley, 2013; Sengupta & Sibley, 2013).

Over time, there may be a kind of “habituation to subjection” (de la Boétie, 1548/2008, p. 54). The disadvantaged may come to tolerate injustice, lower their aspirations, and adapt themselves to unfortunate circumstances (Deutsch, 1985; Major, 1994; Martin, 1986; Moore, 1978). This is especially likely when they view their situation as inescapable or inevitable (Laurin, Kay, & Fitzsimons, 2012; Laurin, Kay, Proudfoot, & Fitzsimons, 2013; Laurin, Shepherd, & Kay, 2010). Even under truly harrowing circumstances—such as slavery or Nazi concentration camps—survivors are known to have made “an adjustment of some sort to the system” so that their “obedience became unquestioning . . . they did not revolt . . . they could not ‘hate’ their masters” (Elkins, 1967, p. 410). On the contrary, many are said to have “identified with the aggressor” (Bettelheim, 1943). No doubt there is an instrumentally rational basis to many acts of obedience; it is often better to comply with powerful others than to risk injury or death. At the same time, it is important—from a social psychological perspective—to distinguish between “excessive” versus “proper meekness,” as Elster (1982, p. 142) put it, or obedience that is “anticipatory” versus compulsory, in the language of Snyder (2017, p. 18).

The point is that rather than criticizing the social system and its authorities (to the extent that criticism is allowable), many people may turn inward and blame themselves and other victims of misfortune. They may, for example, stereotype the poor and exploited as lazy, unintelligent, or otherwise deserving of their plight while lionizing those who are rich and powerful (Allport, 1954; Jost & Banaji, 1994; Kay et al., 2007; Napier, Mändsodza, Andersen, & Jost, 2006). Members of disadvantaged groups may internalize a sense of inferiority, coming to see themselves in less favorable terms than they see members of other, more advantaged groups (Ashburn-Nardo, Knowles, & Monteith, 2003; Clark & Clark, 1947; Jost, Banaji, et al., 2004; Jost, Pelham, & Carvallo, 2002; Lewin, 1941/1948; Rudman, Feinberg, & Fairchild, 2002; Sidanius & Pratto, 1999; Uhlmann, Dasgupta, Elgueta, Greenwald, & Swanson, 2002). As de la Boétie (1548/2008) pointed out, rulers over the ages have cultivated these kinds of outcomes, using “little tricks” to inspire “reverence and admiration” and to sow “doubt in the minds of the rabble as to whether they were not in some way more than man” (p. 66). The Roman emperor Caligula, for example, dressed like a deity, ordered statues to be made with his own head placed on the torsos of the gods, and built a temple in which his subjects could worship him (Kristof, 2017; see also Veyne, 1976/1992). In early 2018, U.S. President Donald Trump sought to inspire admiration and establish authority by taking to social media to insist that he “would qualify as not smart, but genius” and to boast of his “much bigger and more powerful” nuclear arsenal, in comparison with that of a foreign adversary.

It is not too surprising that the powerful would work diligently to maintain their hegemony, but how are we to understand what de la Boétie (1548/2008) regarded as the “convenient gullibility” of the populace? (p. 66). System justification theory suggests that people are motivated to justify existing social, economic, and political arrangements because doing so serves fundamental epistemic, existential, and relational needs (Jost, Langer, et al., 2017). That is, pledging allegiance to the social order gives people a much-needed (even if illusory) sense of certainty, security, and social belongingness. Conversely, those who dare to provoke conflict or challenge established conventions risk an uncertain future involving harsh punishment and social ostracism (e.g., Bowles & Gintis, 2011, p. 109). In evolutionary terms, fitness is increased when the individual is a member (in good standing) of a social system—perhaps any social system, as long as it provides some measure of certainty, safety, and solidarity. Insofar as these are fundamental needs—or core motives—of the human species (e.g., Fiske, 2010), the individual may choose the “security of living wretchedly” over “the uncertain hope of living as he pleases” (de la Boétie, 1548/2008, p. 44).

This does not mean that everyone is equally disposed to justifying the existing regime. People who—for chronic or temporary reasons—are especially eager to attain subjective states of certainty, closure, safety, security, conformity, and affiliation are especially likely to accept and rationalize the “way things are” and to embrace what modern scholars would recognize as politically conservative ways of thinking (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Altemeyer, 1981; Eidelman, Crandall, Goodman, & Blanchar, 2012;
Friesen, Kay, Eibach, & Galinsky, 2014; Jost, 2017a, 2017b; Lammers & Proulx, 2013; Rutjens & Loseman, 2010; Sidanius & Pratto, 1999; Thórisdóttir & Jost, 2011; Ulrich & Cohrs, 2007; G. D. Wilson, 1973). In contrast, individuals who enjoy thinking in complex terms, who are less afraid of death than others, and who value uniqueness over conformity are more likely to criticize the social system and to approve of insurgent movements aimed at changing the status quo (Hennes, Nam, Stern, & Jost, 2012; Jost, Becker, et al., 2017; Jost, Langer, et al., 2017). Thus, in addition to a general tendency for people to adapt themselves to unwelcome realities (T. D. Wilson & Gilbert, 2005), there are individual differences in personality as well as situational triggers pertaining to epistemic, existential, and relational motives that increase or decrease the likelihood of participating in system-challenging collective action (Jost & van der Toorn, 2012; Kay & Friesen, 2011).

It is noteworthy that de la Boëtie (1548/2008) regarded the human proclivity to submit willingly to “the yoke” as akin to the behavior of other beasts of burden. He wrote, “Men are like handsome race horses who first bite the bit and later like it, and rearing under the saddle a while soon learn to enjoy displaying their harness and prance proudly beneath their trappings” (p. 59). Although this passage should be interpreted in a metaphorical sense, because of the pre-Darwin era in which it was written, contemporary researchers may indeed ask whether the toleration of inequality in humans resembles the dynamics of dominance and submission that play out in other species and, indeed, whether it has some basis in biological as well as cultural evolution (e.g., Boehm, 1999; de Waal, 1982; Flannery & Marcus, 2012; Harari, 2015; Hatemi & McDermott, 2011; Kurzban & Leary, 2001; Marcus, 2008; Somit & Peterson, 1997; Sidanius & Pratto, 1999; Tuschman, 2013; Wright, 1994).

An Evolutionary Approach to System Justification Theory

In the remainder of this article, we explore some of the biological bases of system justification in the creative, interdisciplinary spirit of Sapolsky’s (2017) book, Behave: The Biology of Humans at Our Best and Worst. Specifically, we consider a number of parallels between the social behavior of humans and other animals, especially when it comes to the acceptance versus rejection of hierarchy. In a phylogenetic sense, then, we ponder the remarkable capacity for human beings to tolerate oppression and to participate in “voluntary servitude.” Next, we identify neural structures and functions that may, in some sense, undergird ideological processes such as the rationalization of injustice. In sum, we contemplate the evolutionary origins and foundations of system justification tendencies.

Our treatment will be necessarily speculative, because, in the absence of fossil records and direct evidence of structural adaptation under circumstances of natural selection, we are confined to “inferring historical processes from contemporary products” (Richardson, 2007, p. 41). In evolutionary biology, claims about natural selection are evaluated rigorously on the basis of established evidence concerning heritability and genotypic variation at the level of populations. Because our prehistoric ancestors were in no position to leave behind a permanent record of their social behavior, these resources and opportunities are typically not available to evolutionary social psychologists (Buller, 2006; Kitcher, 1985). Indeed, we know very little about the social—structural conditions under which our evolutionary ancestors operated (Richardson, 2007). Consequently, a “recurring problem in applying evolutionary principles to understanding social cognition is that such interpretations are usually post hoc, and notoriously hard to prove” (Forgas, 2007, p. 121). Even our knowledge about the evolution of the human brain is somewhat indirect, insofar as it is based on what can be surmised from the structural features of skulls (rather than brain tissues) recovered through archaeological digs.

Still there may be some theoretical (or meta-theoretical) utility in developing insights—however, speculative—concerning the evolutionary origins of system justification. This is because “an evolutionary perspective helps us to realize that the phenomena we study have biological roots, and offers an important and productive link between cognitive theorizing and the neurosciences” (Forgas, 2007, p. 122; see also Hatemi & McDermott, 2011; Ketelaar & Ellis, 2000; Tooby & Cosmides, 1992). We should ask, for example, why human beings would have evolved the propensity to “love, or at least accept, systems of government that profoundly threaten their self-interest” (Marcus, 2008, pp. 50–51). There may have been certain advantages in evolutionary history—at the aggregate or collective level—to tolerating a rather high degree of dominance, and even oppression—as well as obvious disadvantages to individual members of the species. It is easy to imagine that it would be adaptive for people to be highly attuned to potential threats in the environment, including external threats to the stability of the social system. At the same time, the ability to respond flexibly and openly to changes in the social environment, without undue fear, would also be useful and adaptive.

It is quite likely, in any case, that system justification would contribute to social cohesion (and social order), insofar as quiescence in response to inequity in the social system would signal one’s commitment to social harmony. Much as religious rituals appear to have lent stability to socially stratified societies (Watts et al., 2016), system-justifying ways of thinking may have fostered a sense of legitimacy and therefore popular consent. This notion is consistent with contemporary accounts of the mechanisms of cultural evolution (e.g., Flannery & Marcus, 2012; Norenzayan et al., 2016; Richerson et al., 2016). Coordinated forms of system justification that involve some kind of ideological collaboration between dominant and subordinate members may have increased fitness and survival at the level of the social or cultural system—in comparison with systems that were racked with internal conflict. Of course, any benefits accruing to members of the collective would come at the expense of the health and safety of individual subordinates, as in other species.1
An account of system justification processes in terms of cultural group selection may be the most promising evolutionary model, insofar as it assumes that traits “with some detrimental individual-level effect on reproduction and/or survival can... be favored if group benefits are sufficiently large” (Richerson et al., 2016, p. 8). The idea is that through social psychological mechanisms such as imitation, conformity, and prestige suggestion, human groups differ from one another substantially in terms of cultural norms (e.g., see Flannery & Marcus, 2012), and the cultural norms transmitted by some groups provide a competitive advantage over the norms transmitted by other groups (e.g., Campbell, 1965). Thus, groups in which at least some members participated in “voluntary subordination” (or “anticipatory obedience”) may have been more evolutionarily successful than groups with less social stability in which members were quick to rebel against authorities.

We follow social dominance theory in proposing that there are also meaningful individual differences in political (and prepolitical) preferences for (and against) the maintenance of hierarchical social orders (Sidanius & Pratto, 1999). It is also possible that, at the level of individual fitness, the epistemic, existential, and relational needs to achieve certainty, security, and social belongingness that are thought to underpin system justification motivation are biologically adaptive. That is, it may have been advantageous from an evolutionary standpoint for human beings to develop ways of coping with—and perhaps preventing—feelings of uncertainty, threat, and social isolation. The perspective we are outlining here is most compatible with the assumption that natural selection occurs at multiple levels of analysis, including individual (genetic) and cultural (group) levels (e.g., Sober & Wilson, 1998; D. S. Wilson & Wilson, 2007).

**Animal Models of System Justification**

Not all social species organize themselves in a hierarchical fashion, but a great many do (Sapolsky, 2017). The term “pecking order” comes from notorious browns among hens to establish and enforce “authoritarian” patterns of dominance and subordination (Barth, 2016), but insects and many other species exhibit similar behavior. When two bees are placed together in a space that is too small for more than one nest, a species exhibit similar behavior. When two bees are placed together in a space that is too small for more than one nest, a "queen," stays at the nest and reproduces and guards the nest, while the subordinate female, the ‘worker,’ forages for food” (E. O. Wilson, 2012, p. 142).

In colonies of ants, bees, and wasps, there are fixed hierarchies, caste systems of soldiers and workers, and perhaps even the insect equivalent of “slavery” (E. O. Wilson, 1975). It is common for “workers” to attack one another to maintain the queen’s hegemony. When a female worker attempts to reproduce, “other workers typically thwart the usurper, thus protecting the queen’s primacy”; confederates may, for instance, “pile on the offender to punish her, perhaps severely enough to cripple or kill her” (E. O. Wilson, 2012, p. 145). These dominance hierarchies may be remarkably stable over time, resembling in some ways the successful tyrannical regimes lamented by de la Boétie (1548/2008): “Species organized by despotisms, such as bumblebees, paper wasps, hornets, and artificially crowded territorial fish and lizards... live in relative peace owing to the generally acknowledged power of the tyrant” (E. O. Wilson, 1975, p. 287).

Most, but not all, species of mammals organize themselves into hierarchical social systems, thereby instituting unequal access to resources, such as food, sex, and social status. Hierarchies, according to Sapolsky (2017), “establish a status quo by ritualizing inequalities” (p. 426). To take just one example, a society of naked mole rats in East Africa “occupies and defends a system of subterranean burrows” and enforces a clear division of labor involving a queen, workers, and soldiers, the latter of which risk their lives in defense of the nest (E. O. Wilson, 2012, pp. 41–42). There are clearly material benefits accruing to dominant members of a hierarchical social system, and this explains why they would seek to maintain such a system. In many mammalian species, subordinate members also perpetuate the social order, actively deferring to those above them in the hierarchy (Bernstein, 1969; Zajonc, 1969). Acquiescence is logical when one considers the presumed function of social hierarchies—which is to settle tensions over contested resources with symbols of status differences—rather than actual fighting. It is more adaptive for a diminutive male baboon to relinquish a food item after an alpha male displays his canines threateningly, rather than having to relinquish the item after being bitten by those same canines (Sapolsky, 2017).

In some cases, social rank may be inherited, whereas in others it is based on the results of physical confrontation. For male rhesus monkeys and baboons, high status is generally attained by individuals who are muscular and large in physical size and those having sharp canines and fighting skills. At the same time, maintaining high status has relatively little to do with these factors. Rather, status maintenance is based on social intelligence and impulse control—knowing what others are doing, which coalitions are worth forming, and which provocations should be ignored as well as having the ability to control levels of frustration and aggression to maintain the allegiances of others (Sapolsky, 2017).

Some species are more egalitarian than others. It may be telling that some mammalian species that are most similar to ours, such as male chimpanzees, are clearly hierarchically organized, whereas bonobos, who are as closely related to humans, are much less so. In some cases, there may be a wide gulf between the “alpha” and other members of the group, who are regarded as similar to one another in terms of rank. Or there may be fine—even linear—distinctions in terms of rank, so that each individual understands precisely where he or she is positioned in relation to every other member of the group. Species as diverse as ravens and baboons and human beings are highly attuned to subtle gradations in status and power—and especially to violations or reversals of the dominance hierarchy (Sapolsky, 2017, p. 428). Baboons (and humans) interact quite
differently not only with another individual who is one step above (vs. below) him or her in the hierarchy but also with an individual of an adjacent rank versus one many steps above or below. Con specifics who live in large, socially complex groups (and therefore must track a large number of dominant–subordinate relationships) develop thicker brain regions in some areas, including the prefrontal cortex (Dunbar, 2009; Sallet et al., 2011).

As E. O. Wilson (2012) points out, the “social world of each modern human is not a single tribe, but rather a system of interlocking tribes” (p. 57). Human beings belong to multiple hierarchies that require rather different skills to navigate—and they may occupy very different ranks in each of these hierarchical systems. According to Sapolsky (2017), “this invites rationalization and system justification” (p. 431), insofar as people strive to make sense of relative positions (their own and others’) in the context of multiple, partially overlapping social systems. Harari (2015) writes:

Time and again people have created order in their societies by classifying the population into imagined categories, such as superiors, commoners, and slaves; whites and blacks; patricians and plebeians; Brahmans and Shudras; or rich and poor. These categories have regulated relations between millions of humans by making some people legally, politically or socially superior to others. (p. 136)

Human beings do not behave mindlessly in accordance with relations of dominance and submission. Rather, we use beliefs and ideas to justify why some individuals or groups deserve more resources than others, including material resources necessary for survival as well as symbolic resources such as social status or attention.

These justifications, in turn, represent the cognitive–motivational seeds of political norms or ideologies—defined as socially shared belief systems that either legitimize or delegitimize a given social order (e.g., Hibbing, Smith, & Alford, 2014; Jost, Federico, & Napier, 2009; Tuschman, 2013). When, in ancient times, the Athenians prepared to exploit the Melians, they offered religious justification—“the fate given by the gods to men”—namely, that the “powerful exact what they can, and the weak grant what they must” (E. O. Wilson, 2012, p. 65). The rationalization of inequality on the basis of economic ideology fosters entrenched disparities between rich and poor in contemporary human societies that, according to Sapolsky (2017), are so extreme and long-lasting (even across generations) that they have no parallel in the animal kingdom: “In terms of its caustic, scarring impact on minds and bodies, nothing in the history of animals being crappy to one another about status differences comes within light-years of our invention of poverty” (p. 476). And yet, as Harari (2015) points out, “Most people claim that their social hierarchy is natural and just, while those of other societies are based on false and ridiculous criteria” (p. 136).

One of the many characteristics that sets us apart from other primates, then, is a remarkable capacity to rationalize exploitation and oppression through stereotypes, ideologies, and other belief systems (Jost, Fitzsimons, et al., 2004). At the same time, the seeds of the human capacity for legitimation and rationalization—as well as dramatic inclinations to dominate and submit (that is, to tolerate and even perpetuate one’s own domination)—may originate in behavioral tendencies that are shared with other primates. It has been suggested, for instance, that capuchin monkeys engage in processes of rationalization, such as the devaluation of nonchosen alternatives (Egan, Bloom, & Santos, 2010; Egan, Santos, & Bloom, 2007), and are subject to decision-making biases in favor of the status quo, such as endowment effects and loss aversion (Brosnan et al., 2007; Chen, Lakshminarayanan, & Santos, 2006; Lakshminarayanan, Chen, & Santos, 2008).

Certain ideologies—such as racist doctrines of ethnic superiority—may be wielded to promote violent aggression and dehumanize members of other social groups (Harari, 2015). Consequently, “any brutality can be justified, at any level, and at any size of the victimized group up to and including race and nation” (E. O. Wilson, 2012, p. 63). To explain the occurrence of genocide and war, evolutionary accounts of human social behavior typically emphasize ethnocentric motives and patterns of in-group favoritism and out-group derogation—what Sapolsky (2017) refers to as “Us/Them-ing.” E. O. Wilson (2012, p. 59) writes:

In its power and universality, the tendency to form groups and then favor in-group members has the earmarks of instinct. It could be argued that in-group bias is conditioned by early training to affiliate with family members and by encouragement to play with neighboring children. But even if such experience does play a role, it would be an example of what psychologists call prepared learning . . . . (p. 59)

The idea, in other words, is that human beings are predisposed, not necessarily at a conscious level of awareness, to favor members of one’s own group and disfavor members of other groups (e.g., Fu et al., 2012).

A well-known exception to the phenomenon of in-group favoritism is the tendency for members of socially disadvantaged groups to favor members of advantaged out-groups over members of their own group (Jost & Banaji, 1994; Jost, Banaji, et al., 2004; Sidanius & Pratto, 1999). That is, members of low status or disadvantaged groups—including the poor, stigmatized racial minorities, and gay people—often exhibit out-group favoritism on implicit as well as explicit measures of intergroup bias (e.g., Horwitz & Dovidio, 2015; Jost et al., 2002; Rudman et al., 2002; Uhlmann et al., 2002). This phenomenon, which Lewin (1941/1948) referred to as “group self-hatred,” emerges in early childhood and may persist into adulthood (Baron & Banaji, 2009; Newheiser, Dunham, Merrill, Hoosain, & Olson, 2014). For instance, field studies reveal that Black and Latino passengers tend to tip White taxi drivers better than they tip Black and Latino taxi drivers, respectively (Ayres, Vars, & Zakariya, 2004; Lynn & Withiam, 2008). Out-group favoritism is considered to be a manifestation of system
justification tendencies, and the evidence shows that members of disadvantaged groups who are more politically conservative or score higher on measures of system justification exhibit out-group favoritism to a stronger extent (Ashburn-Nardo et al., 2003; Hoffarth & Jost, 2017; Jost, Banaji, et al., 2004; Pacilli, Taurino, Jost, & van der Toorn, 2011).

Thus far, we have emphasized the ways in which animals—including human animals—accommodate themselves to hierarchical social systems, defer to powerful others, and participate willingly in unequal divisions of labor (see also Harari, 2015). In many species, including species that are closely related to ours, submission in response to dominance behavior is a regular occurrence, and it is conceivable that something analogous occurs in human beings when they are confronted with tyrannical leaders and hegemonic regimes. At the same time, it is important to point out that many species—including ravens, dogs, monkeys, and chimpanzees—are capable of mounting intense protests when they perceive that they are being treated unjustly (Brosnan, 2006). Thus, although they may tolerate inequalities associated with dominance hierarchies, individuals often harbor preferences for equity, equality, and cooperation when it comes to the distribution of food and other resources (Bowles & Gintis, 2011; Brosnan & de Waal, 2003; Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; de Waal, 2009; Tricomi, Rangel, Camerer, & O’Doherty, 2010; Warneken, Lohse, Melis, & Tomasello, 2011). Boehm et al. (1993) have suggested that disobedience is part of human nature, insofar as people often band together to stop exploitative behavior and enforce egalitarian standards. It seems likely that individual differences in psychological characteristics—such as authoritarianism, conservatism, social dominance orientation, and system justification—may help to explain variability in acceptance versus rejection of inequality and other putative injustices (Brosnan et al., 2015; Jost, Glaser, Kruglanski, & Sulloway, 2003; Nam, Jost, Kaggen, Campbell-Meiklejohn, & Van Bavel, 2018; Pratto, Sidanius, Stallworth, & Malle, 1994).

According to Hatemi and McDermott (2011), “Research on human political behavior will be incomplete until it takes into account the evolutionary, neurological, physiological, and genetic foundations of human traits that are universal along with those that vary across groups and individuals” (p. 19). In this section, we have sought to identify fairly universal tendencies that may have arisen through natural selection and that may help to explain the pervasiveness of deference, submission, and tolerance of inequality (Harari, 2015)—as well as the basis for egalitarian revolt (Boehm, 1999). In the next section, we consider variability in such tendencies, focusing on individual differences in system justification and the ways in which these may be instantiated in the human brain.

**Individual Differences in System Justification and the Human Brain**

In general, political scientists tend to assume that people hold the political attitudes they do because of the influence of socialization coming from “top-down” sources, such as the mass media and other elite forms of communication (Fiorina, 2005; Sniderman & Bullock, 2004; Zaller, 1992). Recently, political psychologists have made some headway in identifying “bottom-up” psychological (and even biological) influences or constraints on political attitudes and behaviors (e.g., Ahn et al., 2014; Funk et al., 2013; Hibbing et al., 2014; Jost et al., 2009; Oskarsson et al., 2015). Presumably, differences in ideological orientations linked to system justification processes—such as acceptance versus rejection of inequality and advocacy versus resistance to system change—reflect some degree of genetic influence (Alford, Funk, & Hibbing, 2005; Kandler, Bleidorn, & Riemann, 2012) and are instantiated at the level of specific neuroanatomical structures and functions (Jost, Nam, Amodio, & Van Bavel, 2014).

Amodio, Jost, Master, and Yee (2007) drew on decades of psychological research suggesting that liberals (or leftists) and conservatives (or rightists) differ in terms of cognitive style (Jost et al., 2003; Jost, 2017a) and hypothesized that there would be ideological differences in conflict monitoring, which is a neurocognitive process localized in the anterior cingulate cortex (ACC) that is sensitive to potentially discrepant response tendencies. To test this hypothesis, they administered a “go/no-go” task in which participants are required to respond quickly and accurately to a familiar (go) stimulus, so that “go” responses become habitual. Occasionally, a “no-go” stimulus would appear, and on these trials participants were instructed to withhold their habitual responses. As hypothesized, liberals were more successful than conservatives at responding flexibly by monitoring and inhibiting dominant responses on no-go trials. Importantly, they also exhibited more neural activity in the anterior cingulate during these trials, suggesting that the brains of liberals were more attuned than those of conservatives to the potential for response conflict. Weissflog, Choma, Dywan, van Noordt, and Segalowitz (2013) replicated the results of this study, demonstrating that political liberalism, rejection of inequality, and lower scores on right-wing authoritarianism were all associated with greater anterior cingulate activity on no-go trials. These findings are consistent not only with the behavioral observation that there are ideological differences in self-reported cognitive processing style but also that liberals may be more sensitive than conservatives when it comes to processing novel, unexpected, and ambiguous information (see also Shook & Fazio, 2009).

To investigate the hypothesis, also suggested by Jost, Glaser, Kruglanski, and Sulloway (2003), that there would be ideological differences in physiological responses to potentially threatening stimuli, Oxley et al. (2008) exposed residents of Lincoln, Nebraska, to disgusting and fear-inducing images (such as filthy toilets and gory wounds) as well as neutral images (such as fruit). Results revealed that more socially conservative participants—those who were highly patriotic, nationalistic, religious, and supportive of military spending and “law and order” as well as more traditional lifestyles—exhibited stronger skin conductance responses (increased sweat gland activity associated with arousal in the sympathetic nervous
system) in response to disgusting and threatening, but not neutral, images. In addition, the researchers administered unexpected blasts of white noise and measured the startle eyeblink response, which is considered to index activation in the amygdala, a brain region that is central to the processing of fear and anxiety. Consistent with the results for skin conductance responses, the researchers observed that social conservatives exhibited stronger eyeblink amplitudes, in comparison with social liberals (see also Smith, Oxley, Hibbing, Alford, & Hibbing, 2011).

Investigating the neuroanatomical structure—as well as functional activation—of specific brain regions may provide a useful index of relatively stable individual differences in ideological preferences. Specifically, studies of neural structure assess slow-to-change gray matter volume, which comprises cortical thickness and surface area. Gray matter volume is understood as the computational capacity of a certain brain region, with many studies linking larger gray matter volume with greater efficacy of behaviors supported by that region (Kanai & Rees, 2011). Thus, a landmark study by Kanai, Feilden, Firth, and Rees (2011) addressed the correlation between political orientation and neuroanatomical structure. They measured regional brain volume in university students in the United Kingdom and discovered that more liberal students possessed more gray matter volume in the anterior cingulate, whereas conservative students possessed more gray matter volume in the right amygdala. Given that the ACC is linked to conflict monitoring and the amygdala is centrally involved in threat responses, these differences in neuroanatomical structure fit with theory and evidence suggesting that liberals and conservatives differ with respect to fundamental cognitive–motivational orientations concerning uncertainty and threat (e.g., Jost et al., 2003, 2009). The authors speculate that “it is conceivable that individuals with a larger ACC have a higher capacity to tolerate uncertainty and conflicts, allowing them to accept more liberal views” (p. 678).

To deepen the scientific understanding of ideological differences in neural structure as well as function, Nam, Jost, Kaggen, Campbell-Meiklejohn, and Van Bavel (2018) explored the relationship between system justification and regional brain volume. In addition to the findings of Kanai et al. (2011), there were several reasons to focus on the amygdala in particular. Prior research has established that the amygdala plays a central role when it comes to processing motivationally salient information, including information about threat (Adolphs, Tranel, Damasio, & Damasio, 1995; Phelps et al., 2001) and uncertainty (Herry et al., 2007). In addition, the amygdala appears to be heavily involved in monitoring and facilitating social interaction in group contexts (Van Bavel, Packer, & Cunningham, 2008; Zink et al., 2008). For instance, Kumaran, Melo, and Duzel (2012) observed that larger bilateral gray matter volume in the amygdala was associated with better performance on a task in which participants were required to learn the relative ranks of members in a novel hierarchical social system. (Amygdala volume was unrelated to performance on a learning task involving a nonsocial hierarchy.) Studies also show that the amygdala is especially active in social circumstances in which the individual’s rank is ambiguous and/or unstable (Zink et al., 2008).

Nam et al. (2018) administered Kay and Jost’s (2003) general system justification scale, which measures the belief that the existing social system is fair, legitimate, and operates as it should, to two samples of American adults and conducted structural brain scans. The researchers discovered that individuals who scored higher on system justification possessed more gray matter volume in the bilateral amygdalae (see Figure 1) as well as in parts of neighboring areas such as the insula and orbitofrontal cortex. An analysis comparing relatively high-status individuals (men) with relatively low-status individuals (women) revealed that the relationship between amygdala volume and system justification was statistically indistinguishable for the two groups. This suggests that system-justifying processes may operate similarly for members of dominant and subordinate groups. These findings provide additional evidence that ideology may be reflected in specific brain structures and functions and that the amygdala is a neural substrate that may be centrally involved in maintenance of the societal status quo—for relatively high- and low-status members alike. Strikingly, Nam and colleagues found (in a subset of the initial sample) that having more amygdala volume was associated with a decreased likelihood of participating in system-challenging protests over the subsequent 3 years, strengthening the notion that the amygdala plays a critical motivational role when it comes to defense of the societal status quo.

There is an important caveat, however, that applies to all of the studies that have been conducted thus far in political neuroscience, namely that the direction of causality is ambiguous. We have referred to this conundrum as the “chicken-and-egg” problem in political neuroscience (Jost, Nam, et al., 2014; Jost, Noorbaloochi, & Van Bavel, 2014). Differences in brain activity could play a role in the emergence of ideological differences between liberals and conservatives (or low and high system-justifiers), but it is also conceivable that adopting specific ideological stances subsequently affects brain structure and function. Much as learning how to meditate (Hölzel et al., 2011) or drive taxicab routes (Woollett & Maguire, 2011) changes the structures of specific brain regions (and connections among brain regions), it is possible that immersing oneself in an ideological environment (such as Breitbart News or The Nation) could influence the structural organization and functional operation of one’s brain. Because political attitudes and cortical structures are not fully developed in human beings until early adulthood—and indeed, continue to be somewhat malleable throughout adulthood—political neuroscience is in need of studies designed to isolate causal mechanisms by which changes in brain structure or function affect political behavior (and vice versa).

Studies of patients with brain lesions may help to narrow down the range of interpretations, insofar as damage to certain regions is known to impact specific psychological functions (Anderson & Phelps, 2001; Harrison, Hurlemann, & Adolphs, 2015). Research on brain lesions in nonhuman primates suggests...
that the amygdala is especially important when it comes to navigating the complex, hierarchical social systems of rhesus macaques. For example, individuals tend to drop in the social hierarchy after amygdala lesioning (Bauman, Toscano, Mason, Lavenex, & Amaral, 2006; Rosvold, Mirsky, & Pribram, 1954). Their loss of social dominance may be attributable to a diminished capacity to understand the social (and physical) environment; macaques that received bilateral amygdala lesions exhibited decreased fear in response to the presentation of threatening stimuli such as snakes as well as decreased inhibition when interacting with unfamiliar, potentially adversarial conspecifics (Amaral, 2003). Likewise, humans with amygdala lesions—due to injuries or rare neurological disorders—are less inhibited when it comes to approaching low-information, ambiguous stimuli (Harrison et al., 2015) as well as threatening stimuli (Feinstein et al., 2013). The amygdala is also known to play an important role in species-typical social dominance behavior in other nonhuman animals, such as rats (Bunnell, 1966), cats (Fonberg, 1988), and dogs (Fuller, Rosvold, & Pribram, 1957).

Gray matter volume in the amygdala is associated with high social status in macaques (Sallet et al., 2011) and humans (Bickart, Wright, Dautoff, Dickerson, & Barrett, 2011; Kanai, Bahrani, Roylance, & Rees, 2012), gray matter volume in the amygdala is

**Figure 1.** The association between gray matter volume in the bilateral amygdalae and system justification. (a) Multislice coronal heat maps show gray matter volume differences in the bilateral amygdalae correlated with system justification ($t > 3.0$). Higher system justification tendencies were associated with larger gray matter volume in the bilateral amygdalae as indicated by the overlapping region between bilateral amygdala masks (in blue) and system justification statistical map (in orange). (b) Glass brain image of whole brain analysis (coronal cross-section) suggests specificity of system justification effect in regions including the bilateral amygdalae. (c) Higher system justification was positively associated with larger gray matter volume in the bilateral amygdalae. Here, amygdala volume is computed as the average of left and right amygdala volumes; adjusted for age, gender, and overall brain volume; and standardized such that 0 indicates average volume with changes in 1 standard deviation increments. Source: Nam, Jost, Kaggen, Campbell-Meiklejohn, and Van Bavel (2018; Study 2).
correlated with the size of one’s social network; the assumption is that having a larger amygdala may help individuals to navigate more complex social environments. Taken in conjunction, these findings from lesion studies and the measurement of gray matter volume suggest that the amygdala is a critical brain structure when it comes to interacting with conspecifics, maintaining one’s social ranking, and succeeding in a complex, hierarchical social system.

For all of these reasons, we propose that individual differences in the acceptance versus rejection of existing social systems, with their attendant degrees of hierarchy and inequality, may be traced to structural and functional properties of the amygdala (and adjacent brain regions). According to system justification theory, people tolerate and justify existing forms of inequality to the extent that doing so satisfies underlying needs to manage uncertainty, threat, and social discord (Hennes et al., 2012; Jost, Becker, et al., 2017; Jost, Langer, et al., 2017). Consistent with this analysis, a psychological orientation favoring maintenance of the status quo, which requires vigilance to markers of dominance (and potential changes to dominance rankings) as well as an affinity for hierarchical social arrangements, may find a common neural basis in the amygdala (see also Nam et al., 2018).

General Discussion

In this article, we have recounted long-standing historical and philosophical approaches to understanding how and why people engage in “voluntary servitude,” “anticipatory obedience,” and “false consciousness”—coming to tolerate and perpetuate exploitation, injustice, and oppression (de la Boétie, 1548/2008; Rosen, 1996; Snyder, 2017; Stanley, 2015). Although fear of punishment and rational self-interest may play some role in acquiescence to tyrannical regimes, we conclude that these factors alone cannot fully explain the extent of submission in response to social and political dominance (Gramsci, 1971; Hume, 1742/1987; Lukes, 2011; Moore, 1978). In contemporary social psychology, these issues are addressed most directly by system justification theory (Jost & Banaji, 1994), which suggests that people are motivated, often at a nonconscious level of awareness, to legitimize aspects of the societal status quo, including existing inequalities (Jost, Banaji, et al., 2004; Liviatan & Jost, 2014). Presumably, this is because justifying the overarching social system so helps to address epistemic, existential, and relational needs to manage uncertainty, threat, and social discord (Hennes et al., 2012; Jost & Hunyady, 2005; Jost, Becker, et al., 2017; Jost, Langer, et al., 2017). People are especially likely to legitimize the system and its authorities when they feel highly dependent upon it for their survival (van der Toorn et al., 2015; see also Mason, 1971).

In Pursuit of Consilience

This conclusion is in accord with fundamental discoveries in stress biology, namely, that the most important psychological factors that activate the stress response and increase the risk of stress-related disease include the lack of a sense of control, predictability, and outlets for frustration. The first two factors are sufficiently powerful that, in some circumstances, both human and nonhuman animals exhibit activation of the endocrine stress response when unpredictability increases, even in the presence of increasing rewards (Coates & Herbert, 2008; Levine, Coe, & Wiener, 1989). In some sense, then, we may prefer the certainty of our situation, even if it entails living with unpleasantness, over the uncertainty associated with some unknown prospect for improvement (cf. Fox & Tversky, 1995). Or, to quote the family therapist Virginia Satir, “Most people prefer the certainty of misery to the misery of uncertainty.” It is conceivable that this phenomenon may help to explain working class conservatism—cases in which poor people support political parties and candidates whose economic policies disproportionately hurt members of their own social class (e.g., Jost, 2017b). When the poor vote against their own economic interests, it may be due (in part) to the fact they are choosing the stress-reducing circumstances of stability, familiarity, and predictability.

There are a number of parallels between human and nonhuman animals when it comes to the imposition of hierarchical social systems and adaptation to one’s rank or status within such systems (Sapolsky, 2005). These have inspired a number of attempts to understand the social and political dynamics of dominance and submission in terms of evolutionary psychology (de Waal, 1982; Flannery & Marcus, 2012; Somit & Peterson, 1997; Sidanius & Pratto, 1999; Wright, 1994). At the same time, there is also evidence that members of some species, including humans, sometimes do object to egregious forms of inequality and exhibit inequity aversion (Boehm, 1999; Boehm et al., 1993; Brosnan, 2006; Brosnan & de Waal, 2003; Gurr, 1970; Tricomi et al., 2010; Tuschman, 2013). How to reconcile these two conflicting aspects of human nature—acquiescence and protest—is a fundamental priority for system justification theory, which emphasizes both situational and dispositional variability in ideological motives to accept versus reject the existing social order, with its attendant degree of inequality (Jost, Becker, et al., 2017; Jost, Langer, et al., 2017).

This tension is also reflected in social dominance theory, which contrasts “hierarchy-enhancing” and “hierarchy-attenuating” modes of legitimation (Sidanius & Pratto, 1999). Social dominance and system justification perspectives both assume, therefore, that complex societies contain forces of tension and opposition involving the struggle for dominance versus egalitarianism and the acceptance versus rejection of the existing social order. At the same time, the two theories have different (but possibly complementary) objectives. Whereas social dominance theory seeks to explain the pervasiveness of social inequality and oppression in human societies in terms of evolutionary principles, system justification theory seeks to explain cultural inertia and the success of manufactured consent—that is, the persistence of norms and ideologies that defend and justify social systems that may (or may not) be structured fundamentally in terms of hierarchy and dominance. The point is that once a social system (such as
feudalism, slavery, capitalism, patriarchy, or communism) is firmly established, it tends to be self-perpetuating because of social psychological tendencies that people have evolved to adapt—perhaps “excessively,” at least in a manner of speaking—to the status quo.

Research in social and political neuroscience suggests that biological as well as psychological factors are related to political conservatism, resistance to change, acceptance of inequality, and the strength of system justification motivation (Hibbing et al., 2014; Jost, Nam, et al., 2014). Studies reveal, for instance, that conservatives and high system-justifiers have more gray matter volume in the amygdala (and neighboring areas such as the insula) and less in the ACC, in comparison with liberals and low system-justifiers (Kanai et al., 2011; Nam et al., 2018). This does not mean that individual differences in adolescent brain development cause ideological divergence in adulthood; it is also possible that ideological exposure and indoctrination affect brain structures and functions (Jost, Nam, et al., 2014; Jost, Noorbaloochi, et al., 2014). But these empirical results do suggest that certain political attitudes and behaviors may be linked to biological processes that evolved in other species on the basis of natural selection (Nam, Jost, & Feldman, 2017). The finding that neural processes linked to system justification may be localized to the amygdala—an evolutionarily ancient brain region that is involved in detecting threat and other motivationally salient stimuli—provides a useful data point for speculating about the basic psychological functions underlying system justification. Research on nonhuman primates, which we have alluded to above, indicates that the amygdala is critical to (social) survival, especially in the context of social hierarchy. It is conceivable that individuals who are highly sensitive to threat detection and the nuances of social hierarchy are more likely to survive by accommodating themselves to the existing social order—rather than launching a quixotic challenge that is almost sure to increase risks of threat, uncertainty, and social exclusion (see Bowles & Gintis, 2011).

Our general approach is consistent with the scientific goal of consilience, exemplified by E. O. Wilson’s (1998) doctrine that “the social sciences are intrinsically compatible with the natural sciences. The two great branches of learning will benefit to the extent that their modes of causal explanation are made consistent” (p. 205). It may be worth noting here a truism of evolutionary psychology, most established models of reciprocal altruism (e.g., “tit for tat”) and inclusive fitness (“kin selection”) stress benefits and costs accruing to individual organisms and close kin. But, as Flannery and Marcus (2012) point out, “subordination of self-interest fits poorly with the notion that natural selection operates at the individual, rather than the group, level” (p. 561). Only fairly recently have alternative models of cultural group selection gained traction in evolutionary psychology (Laland & Brown, 2011; Richerson et al., 2016), and these seem better suited to explain system justification and related phenomena such as social dominance. This is because “the subordination of self-interest” and both individual-level and societal-level differences in the normative legitimation of inequality (and, for that matter, equality) cannot be properly understood without taking into account mechanisms of cultural as well as biological evolution (e.g., Flannery & Marcus, 2012).

But it is striking to see how heavily the most prominent evolutionary accounts of social and political behavior today depend upon the assumption that behavior is motivated—either consciously or unconsciously—by rational self-interest (Kenrick & Griskevicius, 2013; Petersen, Sznyier, Sell, Cosmides, & Tooby, 2013; Pinsof & Haselton, 2016, 2017; Weeden & Kurzban, 2017). With respect to intergroup relations, this means that they emphasize the universality of in-group favoritism and tend to overlook the occurrence of out-group favoritism (e.g., Jost, Banaji, et al., 2004). For example, Weeden and Kurzban (2014) argue that political preferences, including ideologies, are determined by “inclusive interests” (p. 42), which they characterize as “Machiavellian,” addressing fundamentally selfish concerns: “What’s in it for me, my family, my friends, my allies, and my wider social network?” This framework does a reasonable job of accounting for central tendencies, that is, explaining why poorer people are more supportive than wealthier people of economic welfare policies, women are more supportive than men of gender equality, and African Americans are more supportive than European Americans of race-based affirmative action policies. But even if women and members of the working class are (on average) slightly more liberal than men and members of the middle or upper classes, there are millions of conservative women and millions of working-class conservatives, and their attitudes and behaviors must be explained as well (Jost, 2017b).

In an instructive application of Weeden and Kurzban’s (2014) self-interest theory, Pinsof and Haselton (2016, 2017) proposed that heterosexuals who are more politically conservative would be driven to oppose same-sex marriage because of stereotypical associations linking gay men and lesbians to

**Implications for the Study of Evolutionary Psychology**

In closing, we wish to add a few words about the implications of our work for the study of evolutionary psychology, on one hand, and political science, on the other. With respect to evolutionary psychology, most established models of reciprocal altruism (e.g., “tit for tat”) and inclusive fitness (“kin selection”) stress benefits and costs accruing to individual organisms and close kin. But, as Flannery and Marcus (2012) point out, “subordination of self-interest fits poorly with the notion that natural selection operates at the individual, rather than the group, level” (p. 561). Only fairly recently have alternative models of cultural group selection gained traction in evolutionary psychology (Laland & Brown, 2011; Richerson et al., 2016), and these seem better suited to explain system justification and related phenomena such as social dominance. This is because “the subordination of self-interest” and both individual-level and societal-level differences in the normative legitimation of inequality (and, for that matter, equality) cannot be properly understood without taking into account mechanisms of cultural as well as biological evolution (e.g., Flannery & Marcus, 2012).

But it is striking to see how heavily the most prominent evolutionary accounts of social and political behavior today depend upon the assumption that behavior is motivated—either consciously or unconsciously—by rational self-interest (Kenrick & Griskevicius, 2013; Petersen, Sznyier, Sell, Cosmides, & Tooby, 2013; Pinsof & Haselton, 2016, 2017; Weeden & Kurzban, 2017). With respect to intergroup relations, this means that they emphasize the universality of in-group favoritism and tend to overlook the occurrence of out-group favoritism (e.g., Jost, Banaji, et al., 2004). For example, Weeden and Kurzban (2014) argue that political preferences, including ideologies, are determined by “inclusive interests” (p. 42), which they characterize as “Machiavellian,” addressing fundamentally selfish concerns: “What’s in it for me, my family, my friends, my allies, and my wider social network?” This framework does a reasonable job of accounting for central tendencies, that is, explaining why poorer people are more supportive than wealthier people of economic welfare policies, women are more supportive than men of gender equality, and African Americans are more supportive than European Americans of race-based affirmative action policies. But even if women and members of the working class are (on average) slightly more liberal than men and members of the middle or upper classes, there are millions of conservative women and millions of working-class conservatives, and their attitudes and behaviors must be explained as well (Jost, 2017b).
promiscuity. The idea is that these stereotypes threaten conservative heterosexuals’ interests in pursuing “short-term mating strategies,” that is, marrying in early adulthood, settling down, and raising a family. However, Hoffarth and Jost (2017) challenged the interpretation of the connection between conservative ideology and opposition to same-sex marriage in terms of self-interest motivation. They suggested that ideological factors—such as individual differences in system justification motivation—would explain conservative opposition to same sex-marriage better than self-interest per se (see also van der Toorn, Jost, Packer, Noorbaloochi, & Van Bavel, 2017). Thus, Hoffarth and Jost hypothesized that the same positive association would obtain between conservatism and opposition to same-sex marriage among sexual minorities—for whom the logic of self-interest did not hold. In support of this hypothesis, past research shows that gay men and lesbians who are more (vs. less) politically conservative display more out-group favoritism, expressing implicit as well as explicit preferences for heterosexuals in a number of domains (Jost, Banaji, et al., 2004; Pacilli et al., 2011).

Hoffarth and Jost (2017) reanalyzed data from sexual minorities who had participated in Pinsof and Haselton’s (2016) study and found that, for this subsample, people who were more conservative and who endorsed the homosexuality-promiscuity stereotype both implicitly and explicitly were less supportive of same-sex marriage, compared to those who were more liberal and who rejected the stereotype. The results, in other words, mirrored those obtained for the predominantly heterosexual sample analyzed by Pinsof and Haselton. Because this subsample of sexual minorities was rather small, Hoffarth and Jost analyzed data from two other surveys and observed that in both cases, more (vs. less) conservative sexual minorities were less supportive of same-sex marriage. Thus, ideology may provide a better explanation than self-interest when it comes to understanding the attitudes and behaviors of groups that are disadvantaged or devalued within the context of the overarching social system (see also Crandall, 1994; Jost, Banaji, et al., 2004).

Evolutionary approaches that focus exclusively on self-interest motivation will, in our view, be unable to provide a truly satisfactory account of social and political psychology, especially when it comes to the dynamics of dominance and submission, exploitation and legitimation, and ideology and the justification of inequality. For instance, Norenzayan et al. (2016) provide a sophisticated analysis of the ways in which religious beliefs and practices benefited our evolutionary ancestors by promoting shared interests and prosocial behavior, such as large-scale cooperation among strangers, but their framework has little to say about the antisocial aspects of religion, such as the ideological cover that religious beliefs provide for inequality, injustice, and exploitation (Flannery & Marcus, 2012; Jost et al., 2014; Watts et al., 2016). The fact that certain behaviors or sets of behaviors were biologically adaptive in the ancestral past by no means guarantees that they are beneficial for us today or in the future. System justification in the era of global capitalism and anthropogenic climate change may well spell disaster for the human species (e.g., Klein, 2014).

Implications for the Study of Political Behavior

Although we regard much of what we have written here to be speculative, we agree heartily with Hatemi and McDermott (2011) that “Political behavior scholars can use biology and evolution to construct empirically falsifiable hypotheses and thus expand the discipline’s explanatory capacity” (p. 37). Our primary objective in this article has been to stimulate theoretical development with respect to issues of system justification, especially the question of why people—often, but not always— acquiesce in the face of exploitation and injustice and lend legitimacy to social systems that are hierarchical and putatively oppressive. We have addressed this objective by considering the logic of evolutionary theory as well as biological findings based on animal models of dominance and submission and recent progress in human neuroscience (see also Sapolsky, 2017). If our theoretical (and meta-theoretical) approach is sound, then a number of implications for the scientific study of political behavior follow.

For one thing, system justification theory may help to explain why protest and revolution are such rare occurrences. According to recent data from the World Values Survey, less than one in five citizens from North America, Western Europe, Australia, and New Zealand have participated in a political demonstration, and more than one-third state they would never participate in such a demonstration (Jost, Becker, et al., 2017). Given the extent of social inequality and economic exploitation that persists (e.g., Atkinson, Piketty, & Saez, 2011; Corning, 2011; Frank, 2010; Hacker & Pierson, 2010), the reluctance to protest—especially in Western, democratic nations, where the right to participate in collective action is constitutionally protected—is difficult to square with prominent theories in political sociology that assume (based on principles of rational self-interest) that people are “quick to anger” when confronted with personal and group-based deprivation (Gurr, 1970, p. 58). The perspective we have sought to develop here is far more consistent with the observation of Kinder and Sears (1985) who concluded that “the deepest puzzle here is not occasional protest but pervasive tranquility” (p. 702). System justification theory, it should be clear by now, reflects a social–psychological attempt to understand this puzzle (see also Jost, Gaucher, & Stern, 2015; Jost & van der Toorn, 2012).

There are implications here for international relations as well. As Rothbard (1975/2008) pointed out, the concept of voluntary servitude, as introduced by de la Boétie (1548/2008)—and, by extension, the occurrence of system justification (Jost & Banaji, 1994)—is seldom appreciated by foreign policy experts, who assume that citizens in other countries (a) are motivated more by rational self-interest than they actually are and (b) hold political preferences that are more similar to those observed in the experts’ own countries than they actually are. To take just one salient example from the Cold War era:
[A]nti-Communists write about Communist rule as if it were solely terror imposed from above on the angry and discontented masses. Many of the errors of American foreign policy have stemmed from the idea that the majority of the population of a country can never accept and believe in Communist ideas, which must therefore be imposed by either a small clique or by outside agents from existing Communist countries. (Rothbard, 1975/2008, p. 33)

For instance, the great nuclear physicist, Sakharov (1990), looked back on his own “voluntary servitude” to the Soviet Union as follows:

I felt myself committed to the goal which I assumed was Stalin’s as well... I needed, as anyone might in my circumstances, to create an illusory world, to justify myself... the state, the nation, and the ideals of communism remained intact for me... In the face of all I had seen, I still believed that the Soviet state represented a breakthrough into the future, a prototype (though not yet fully realized one) for all other countries to imitate. That shows the hypnotic power of mass ideology. (p. 64)

Many experts, we suspect, fail to grasp the extraordinary extent to which human beings—perhaps even themselves—are motivated to defend, bolster, and justify the social, economic, and political systems on which they depend (see also Kuran, 1991, p. 31).

Implications for the Study of Social Change

It follows from system justification theory that social change will be difficult for social and psychological—as well as political—reasons, but by no means are we suggesting that protest and resistance are unheard of. Obviously, revolutions occur, albeit infrequently. The literature on system justification suggests that there are several factors that increase the likelihood that people will fight to change the social system (Gaucher & Jost, 2011). First, there are other motives identified by the theory—such as ego and group justification motives to defend and justify individual and collective interests—that may overcome system justification motivation in some situations. When the social system clearly fails to provide a sense of safety, security, and social belongingness, those who are at risk may challenge the system out of a desire for self-preservation. Recent examples include the Black Lives Matter protest movement, which emerged in response to police killings of unarmed African American men, and the #MeToo movement in which women used social media platforms to denounce perpetrators of sexual harassment and sexual assault.

Second, human beings presumably evolved a number of other psychological tendencies—such as motives for accuracy, justice, innovation, and reform—that may likewise trump system justification tendencies. For instance, people commonly exhibit inequity (or inequality) aversion—even when they benefit from the inequity (Fehr & Schmidt, 1999; Rilling & Sanfey, 2011), and nearly everyone has a vested interest in improving upon the status quo (Johnson & Fujita, 2012). Consistent with Lewinian field theory, which assumes that social situations typically involve multiple, potentially conflicting psychological forces, system justification theory recognizes a number of different goals or motives that may be in conflict or tension with one another. For instance, “a person may feel torn between the government and the opposition, seeing both advantages and disadvantages to the existing regime” (Kuran, 1991, p. 17). There is no a priori reason to assume that system justification motivation will necessarily “win out” over other goals or perceptions.

Third, even when system justification motivation is highly salient or activated, there is evidence that people are less defensive and more supportive of social change when it is framed as “system sanctioned”—that is, congruent rather than incongruent with the norms and ideals of the overarching social system. For instance, high system-justifiers were found to be more supportive of proenvironmental initiatives when these were framed as necessary for maintaining the “American way of life” (Feygina, Jost, & Goldsmith, 2010). The success of the civil rights movement under Martin Luther King, Jr., may be attributable, at least in part, by the leader’s skill in appealing often to system-sanctioned ideals, as in his famous speech from 1963:

... when the architects of our republic wrote the magnificent words of the Constitution and the Declaration of Independence, they were signing a promissory note to which every American was to fall heir... that all men would be guaranteed the inalienable rights of life, liberty, and the pursuit of happiness... I have a dream that one day this nation will rise up and live out the true meaning of its creed.

It is possible that system-sanctioned appeals are less threatening than other types of proposed changes, insofar as the backdrop of social stability is largely maintained: The status quo can be improved and modified without being overthrown.

Fourth and finally, research on system justification theory suggests that once a new regime is perceived as inevitable (or nearly so), people should be motivated to engage in anticipatory rationalization of the new status quo (Kay, Jimenez, & Jost, 2002; Laurin et al., 2012). In other words, people may resist change and tolerate the existing social order until a viable alternative system begins to take its place, at which point the dependence on the old system is lessened; it then loses psychological significance. Such a formulation is broadly consistent with “tipping point” models of “revolutionary bandwagons,” such as that of Kuran (1991), who noted that for several decades, “most East Europeans displayed a remarkable tolerance for tyranny and inefficiency. They remained docile, submissive, and even outwardly supportive of the status quo” (p. 26). However, as public opposition to the Communist system spread rapidly across Poland, East Germany, Hungary, Czechoslovakia, and Romania in the late 1980s and early 1990s, revolution seemed imminent, and one nation after another turned “with little warning from quiescence and subservience to turbulence and defiance” (p. 42).
Interestingly, the toppling of hierarchies in some other species likewise follows a nonlinear pattern. In the case of baboons, for instance, there may be a long period in which the top rival of an alpha male looms but is intimidated repeatedly into backing down. Eventually, they fight and the rival wins the confrontation handily. In the ensuing weeks, the ex-alpha drops precipitously in the hierarchy, losing one dyadic struggle after another. In retrospect, it becomes clear that he had been maintaining his position through intimidation and past reputation, but physically he was well past his prime. This situation is such that a stale hierarchy may persist for a very long time, and when change finally occurs, it occurs suddenly and dramatically.

Concluding Remarks

In many ways, a system justification perspective is fairly pessimistic about prospects for human thriving. People appear to be remarkably willing to tolerate exploitation, injustice, and even tyranny and to make virtues of necessity. As Moore (1978) put it, “The human capacity to withstand suffering and abuse is impressive, tragically so” (p. 13). To make ourselves feel better about the present state of affairs—the things we cannot change—we often devise stereotypes, rationalizations, and even full-fledged ideologies to justify the way things are. Better to stick with the “devil we know” than to encounter ones we do not. As a result, the first order of business for antisystem dissidents is to “shatter the appearance of the invulnerability of the status quo” (Kuran, 1991, p. 25). Until then, it is an uphill struggle, to say the least, to push for qualitative social change.

But there is a much more optimistic consequence of system justification theory. It is that once social systems that enshrine liberty, justice, and democratic norms that are protective of diversity, tolerance, and equality are firmly established, people should be motivated to defend and justify these social systems as well. The idea at the core of system justification theory is that people are motivated to preserve the social, economic, and political arrangements that they inherit from their cultural predecessors—arrangements that are comfortable to them and in which stressors are familiar and predictable (Jost, 2015). After all, these are the institutions to which they have pledged allegiance since childhood (Eibach, Wilmot, & Libby, 2015; Jost, Sterling, & Langer, 2015). “The things we do,” as Festinger (1983) pointed out, “are always good” (p. 163). In principle, at least, our institutions need not be exploitative, oppressive, or unjust. Whereas social dominance theory holds that egalitarianism could never provide the basis for a truly stable equilibrium in human social systems (Sidanius & Pratto, 1999), system justification theory does not.

That is, it is the established order that, from a system justification perspective, “reproduces ideology in the minds of successive generations” (Veyne, 1976/1992, p. 379). By historical fiat, the institutional (and ideological) inheritance may be either fortunate or unfortunate, just or unjust, oppressive or exploitative—or otherwise. Thus, de la Boétie (1548/2008) noted that, “It was impossible for the Persian to regret liberty, not having known it, nor for the Lacedaimonians to find subjection acceptable after having enjoyed freedom” (p. 57). For better and worse, we accommodate ourselves, psychologically speaking, to the social systems on which we depend. Rothbard (1975/2008, p. 36), a devotee of de la Boétie, made a similar point with an eye trained firmly on the precarious future of mankind: “If a free society were ever to be established, then, the chances for its maintaining itself would be excellent.”

Acknowledgments

We thank Rose Mcdermott, Robert C. Richardson, Jim Sidanius, and two anonymous reviewers for providing extremely helpful feedback on an earlier draft of this article. In addition, we acknowledge very constructive suggestions for revision made by Flávio Azevedo, Vivienne Badaan, Riana Brown, Sarah DiMuccio, Shahzad Goudarzi, Mark Hoffarth, Mao Mogami, and Jussi Valtonen. Finally, we are grateful to Michelle Io-Low for assistance with the Reference section, Jaime Napier for recommending Sapiens, Jean E. Jost for recommending On Tyranny and Catalin Mamali for bringing Andrei Sakharov’s Memoirs to our attention.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The lead author was supported in part by National Science Foundation Award # BCS-1627691 during the period in which this article was written.

Notes

1. It is also conceivable that system justification is quite simply an accidental evolutionary by-product—what Marcus (2008) describes as a “kluge”: the unfortunate, perhaps tragic, result of other social and psychological adaptations.

2. Using a more lenient statistical criterion, whole brain analyses conducted by Kanai, Feilden, Firth, and Rees (2011) revealed that conservatives also had more gray matter volume in the left insula and right entorhinal cortex in comparison with liberals.

3. Other research suggests that “internalized homonegativity” (an analogue of out-group favoritism or “group self-hatred”) among gay men and lesbians is associated with increased stress, elevated cortisol levels, and depression (Bahamondes-Correa, 2016; Lewis, Derlega, Griffin, & Krowinski, 2003; Parra, Benibgui, Helm, & Hastings, 2016).

References


Elster, J. (1982). Belief, bias, and ideology. In M. Hollis & S. Lukes
Godfrey, E. B., & Wolf, S. (2016). Developing critical consciousness or justifying the system? A qualitative analysis of attributions for poverty and wealth among low-income racial/ethnic minority and


