Why Be Immoral?[1]

Christopher Freiman
College of William and Mary
cafreiman@wm.edu

No one chooses to have everything if he has first to become someone else.
Aristotle[2]

The question "Why be moral?" is so central to moral theory that H.A. Pritchard dubbed it *the* moral question (1912).[3] But the complementary question, "Why be immoral?" has received relatively no attention.[4] Why?

Here is one possibility. We assume we already know the answer: self-interest. Consider the paradigm immoralists. Callicles rejects morality to satisfy his desires. Hobbes's Foole defects from cooperation for personal gain. Hume grapples with the sensible knave; Gauthier with the straightforward maximizer. We take it for granted that immorality is animated by an eagerness to defy social norms that restrict our ability to achieve our ends.

This picture of the immoralist, however common, is crucially incomplete. We overlook the possibility, suggested by Kant and Rousseau, that immorality can be animated by an eagerness to conform to social norms that actually sabotage our ability to achieve our ends. What Kant calls "servility" is an underappreciated but pervasive motive for immorality.[5]

Recognizing servility as a basic cause of immorality obliges us to reconsider questions about the rationality of morality. If we are to offer compelling reasons to refrain from immorality, we must offer compelling reasons to refrain from servility. My account can be located within a tradition

---

[2] Aristotle (2005, 1166a).
[3] Christine Korsgaard (2005) calls it "the normative question;" Alan Gewirth (1981), "the authoritative question."
[4] For notable exceptions, see Vogler (2002, p. 5) and Gert (1998).
[5] The motive of servility in some respects resembles Rousseau's notion of *amour-propre*, which plays a central role in his theory of immorality.

that claims a satisfactory answer to the moral question will justify the demands of morality from the first-person perspective.[6] On this conception, to show that it is rational to be moral, we need to show that the balance of reasons favors morality from individuals' own point of view. We must give an answer to the moral question that individuals themselves regard as compelling enough to provide motivation to do what's right.[7]

Accordingly, answering the moral question requires a profile of individuals' motivations. If we are to offer individuals reasons to be moral that they regard as defeating their reasons to be immoral, we need to know what considerations they regard as favoring immorality. An account of the motives for immoral behavior constrains answers to the moral question.

Consider, for example, that philosophers as diverse as Plato, Hobbes, Hume, and Gauthier are unanimous in their agreement that (misperceived) self-interest drives immorality. On this view, an adequate answer to the moral question—an answer that we can accept from our own point of view—must then appeal to our self-interest. For this reason, these philosophers' remedies for immorality amount to demonstrating, in some way, the concordance of morality and *rational* self-interest. Self-interested individuals will accept the authority of morality only upon being shown that true profit consists in living virtuously.

---

[6] See Korsgaard (2005).

[7] There is disagreement about the relationship between motivating reasons and normative reasons. According to reasons internalism, having an available motivating reason to φ is necessary for having a normative reason to φ. Reasons externalism, in contrast, denies that normative reasons must be capable of motivating us. For a recent discussion, see Shafer-Landau (2003, p. 165-189). Externalists will therefore deny my assumption that showing that the balance of reasons favors morality from the point of view of agents themselves is necessary to showing that it is rational to be moral. Defending reasons internalism is beyond the scope of this paper; however, as Korsgaard notes, philosophers' attempts to answer the moral question are often conducted on the assumption that an answer must address the first-person position—and thus the motivational commitments and propensities—of the agent seeking a justification for morality. Korsgaard (2005, p. 49ff.) offers Hume, Bernard Williams, Mill, Kant, and possibly Aristotle as examples. Thus, although some will reject my assumption of reasons internalism, my argument proceeds in accordance with an established approach to answering the moral question. I am grateful to an anonymous reviewer for pressing me to clarify this point.

But suppose we deny that our motivations are predominately egoistic. Indeed, I will argue that we sometimes act immorally at the knowing expense of our personal interests. Traditional answers to the problem of the immoralist will be of little help here. To furnish those individuals led to immorality by servility with a compelling reason to be moral, we must appeal to some justification other than self-interest.

In response to the challenges presented by servile immorality, I will field a new argument on behalf of the rationality of morality. At first blush, this argument resembles what David Copp (2004) calls the "self-conception strategy." This style of argument is typically, if not exclusively, affiliated with Kantians. The basic idea is that agents' conceptions of themselves supply them with reason to comply with morality's demands. Korsgaard, for instance, claims that to violate morality "is to no longer be able to think of yourself under the description under which you value yourself and find your life to be worth living . . . It is to be for all practical purposes dead or worse than dead" (2005, p. 102). It is rational to be moral because doing the wrong thing amounts to "not being ourselves any more" (2005, p. 18).

But many doubt that the self-conception strategy can deliver on its rhetorical promise. We should be skeptical of claims implying that accepting a bribe leaves one "for all practical purposes dead or worse than dead," or that breaking a promise necessarily subverts one's rational agency, one's humanity. Moreover, even if prevailing Kantian formulations of the strategy successfully show that moral claims issue from our humanity, they may still fall short of imbuing morality with the overriding force Kantians ascribe to it.

The worry can be expressed in terms of a familiar objection to Kantian ethics—its conception of the self is too "thin." Korsgaard claims that obedience to morality is obedience to our humanity, shorn of our particular interests, desires, and projects. But why we, thick with

contingences, should be heartened that by acting morally we act in accordance with the demands of the rational being within us—when doing so can cost our contingent selves dearly—is unclear.[8]

What *would* imbue morality with the force Korsgaard ascribes to it? Suppose that transgressing was analogous to, say, wearing a golden ring one stumbled upon while shepherding. Turning the collet of the ring brings small comforts and conveniences. But there's a catch: the ring's magical power has the unfortunate side effect of radically altering your identity. Each turn of the collet causes a corresponding change in your personality. Over time, you won't recognize the people you now love. You won't care about the projects that now give your life meaning. You won't hold the beliefs you now cherish.

The choice to wear the ring is, in a meaningful sense, suicidal: it is to be for all practical purposes dead or worse than dead. Your ends, attitudes, and projects are part of what makes you, you. But what's so nightmarish about wearing the self-deleting ring isn't the prospect of losing your humanity (you wouldn't); rather it's the prospect of losing the contingent dimensions of your identity—those particular concerns that give your life meaning.

Intuitively, you have a reason not to use the ring (perhaps not even once). Certainly using the ring would make your life easier, but the thought is that it would no longer be *your* life. Kantians may be unable to accommodate this judgment, as using the ring would not bring about a loss of our humanity. Korsgaard (2005, p. 102) asserts that our moral obligations proceed from "the conceptions of ourselves that are most important to us." But the conceptions of ourselves that are most important to us need not be those conceptions of ourselves that are necessary. Most of us endorse and closely identify with aspects of our identity that can be abandoned without thereby abandoning our humanity.

---

[8] For a similar criticism and formulation, see Nozick (1974, p. 228)

If immorality was like using the ring, our reason to resist immorality would be clear. This paper is an attempt to show that sometimes immorality *can* be like using the ring. Servility is characteristically motivated by an aversion to the discomfort attendant on resisting social pressure; however, servility can lead to an undesired transformation or dissolution of those particular desires, beliefs, and projects that constitute agents' most highly valued contingent conceptions of themselves. Thus, my argument resembles a Humean variant of the self-conception strategy: agents have reason to abstain from servility in order to preserve the integrity of their "thick" practical identities.

In (§1), I argue that evidence concerning the motives of immoral behavior belies the assumption that immorality is predominately motivated by self-interest. (§2) takes up Thomas Hill's suggestion that servility disposes individuals to immorality, rendering them poorly positioned to recognize and affirm a proper respect for others. In (§3), I argue that servile behavior implies cognitive dissonance, which can restructure an agent's personality over time. (§4) deploys this conclusion to establish that we have reason to abstain from servility, even on a parsimonious Humean account of practical reasons. (§5) discusses the role of habituation in substantiating the rationality of morality. (§6) concludes.

§1

The claim that self-interest cannot serve as a comprehensive explanation of immorality finds support in ordinary moral experience, historical example, and several celebrated findings in social psychology. By way of example, consider the following:

DUPLICITOUS NEWCOMER: A Whig moves to a Tory community (or vice versa). She finds herself surrounded by people with opposing political views—at work, at school, and in the pubs.

Whenever people start discussing politics, she lies. She voices agreement with their views.[9] She lies, in spite of knowing she may thereby dissuade people from her ideals.

ABU GHRAIB: In 2004, it was discovered that American soldiers physically and psychologically tortured prisoners of war at Abu Ghraib. Detainees were punched, slapped, and kicked. They were sexually abused. Some were threatened with rape.[10] Although the reasons behind the soldiers' behavior were complex, social psychologist Susan Fiske and colleagues concluded that conformity to peers and socialized obedience to authority were major factors. Soldiers felt pressure to conform to the "perceived reactions" of the group and to obey the actual orders of superiors as well as the orders they *thought* their superiors would issue (Fiske et al. 2004, p. 1483).

RESERVE BATTALION 101: In 1942, Reserve Battalion 101, a unit of roughly 500 men from Hamburg, Germany were recruited to exterminate all Jews living in the remote towns of Poland.[11] The recruits were sent to Poland without any knowledge of their mission. They were told their mission's purpose only upon arrival in Poland. In four months, they shot 38,000 Jews at point blank range and deported 45,000 more to a concentration camp in Treblinka.

The battalion's commander told recruits that they could refuse to execute these people. Indeed, at first, about half of the recruits did refuse. "But over time," social psychologist Philip Zimbardo reports, "social modeling processes took over, as did guilt-induced persuasion by those reservists who had been doing the shooting, along with the usual group conformity pressures of 'how would they be seen in the eyes of their comrades'" (Zimbardo 2007, p. 286). At the end, 90 percent of those in Battalion 101 obeyed their battalion leader and became personally involved in the executions.

---

[9] For a classic study of conformity of normative opinion, see Crutchfield (1955).
[10] For an overview, see Zimbardo (2007).
[11] For discussion, see Browning (1992).

Similar obedience and conformity effects have been replicated in familiar psychological experiments. In Zimbardo's own Stanford Prison Experiment (which has drawn explicit comparisons to Reserve Battalion 101 and Abu Ghraib) psychologically normal individuals were recruited to study the psychological effects of prison life.[12] Participants were arbitrarily divided into "guards" and "prisoners" to serve in an artificial prison at Stanford University. The guards were pressured to keep order among the prisoners at all costs (outside of violence).[13] Guards that failed to use their power were ridiculed by their fellow guards for their failure to conform to norms of guard behavior. All of the guards behaved sadistically toward the prisoners at some point.[14] Many guards acted cruelly toward the prisoners, not because they enjoyed it (although some did), but because they could not bring themselves to violate the "etiquette" of the prison and resist the requirements of their social roles.

Investigating the psychology of the Nazi executioners, Stanley Milgram conducted a series of famous experiments to test obedience. Milgram found that most subjects would obey orders to torture innocent strangers. They administered what appeared to be painful and harmful electric shocks to another human being simply because an authority figure told them to do so. The subjects were "teachers" ordered by the experimenter to shock a "learner" whenever the learner forgot a word pair. Unknown to the subjects, the learner was a confederate receiving no shock. Subjects continued to shock even while hearing the learner scream and demand that the shocks be stopped. Over half of them delivered what they thought was a maximum 450 volt shock to the learner. The

---

[12] For the comparison to Reserve Battalion 101, see Browning (1992, p. 168). For the comparison to Abu Ghraib, see Zimbardo (2007, p. 324-379).

[13] The guards were told "that they must maintain 'law and order' in this prison, that they were responsible for any trouble that might break out" (Zimbardo 1973, p. 40).

[14] Zimbardo writes, "Everyone and everything in the prison was defined by power. To be a guard who did not take advantage of this institutionally sanctioned use of power was to appear 'weak,' 'out of it,' 'wired up by the prisoners,' or simply a deviant from the established norms of appropriate guard behavior . . . [A]ll of the mock guards at one time or another during this study behaved sadistically toward the prisoners" (Zimbardo 1973, p. 49).

majority of subjects stopped shocking only when the experimenter told them to. By that point, the victim was silent, presumably unconscious or worse.[15]

Standard analyses of immorality in terms of self-interest simply do not adequately explain the preceding cases of viciousness. It would be misleading to characterize these acts as self-interested. Zimbardo attributes the murders of Battalion 101 to recruits' inability to violate the expectations of peers. He writes, "[T]here was no special selection of these men, nor self-selection, nor self-interest or careerism that could account for these mass murders" (2007, p. 286).

Milgram's subjects gained nothing by shocking the learner and they faced no penalties for disobedience. They knew they would keep their fee even if they discontinued the experiment. They didn't face a danger and were free to leave at any time. They acted immorally but not from the anticipation of a personal gain. None of them wanted to act as they did, in the sense of advancing their own interests. Milgram concludes that "few people have the resources needed to resist authority. A variety of inhibitions against disobeying authority come into play and successfully keep the person in his place" (1974, p. 6).

Similarly, one guard in Zimbardo's prison wrote that he was "continually called upon to act in a way that was just contrary to what I really feel like inside" (1973, p.49). Nonetheless, he answered that call to act. David Schmidtz (1997, p. 121) writes, "The fact that it can be hard to be moral is, I infer, not always self-interest's fault. Doing the right thing in the face of social pressure to do otherwise is not easy." The individuals in the above examples are best described not as pursuing interests of the self, but denying them.[16] What they have in common is an unwillingness to assert themselves, especially when doing so violates the preferences or expectations of others.

One might deny that these individuals are not self-interested—perhaps they simply desire to conform more than they desire to assert themselves. Of course, desires understood as intentional

---

[15] For a summary of the experiments, see Milgram (1975).
[16] Similar themes are explored in Hampton (1993).

states with a world-to-mind direction of fit need not be self-regarding. Interests *of* the self need not be interests *in* the self. Individuals can have tuistic desires; sometimes we take an interest in others' interests.

We might describe the individuals in the above examples as having a tuistic, *de dicto* desire to do what others expect or demand of them. Moreover, this desire is incompatible with at least one of their *de re* desires, such as not causing needless pain. Perhaps as a result of this conflict, conformity was often accompanied by anxiety rather than the positive affect which is a phenomenological marker of advances in one's own well-being. Milgram reports that his obedient subjects would "sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their flesh" (1963, p. 375).[17] The subjects' conformity entailed subordinating their independently held *de re* desires to their *de dicto* desire to yield to the preferences of others, whatever they might be. The subjects contrast with the paradigm case of the self-interested immoralist, whose immoral behavior is due partly to a tendency to disregard the preferences of others. The immoral behavior of the individuals above is due partly to a tendency to *accommodate* the preferences of others. We can characterize their subordination as *servile* rather than self-serving.[18]

§2

---

[17] Milgram's description recalls Aristotle's (2005, 1166b) claim that those who act badly "are in internal conflict, and have an appetite for one thing but wish for another . . . [T]heir soul is in a state of civil strife, and one element in it, because of its wickedness, grieves in abstaining from certain things, while the other element is pleased; the one draws them this way, the other that, as if tearing them apart." Commenting on the above passage, Jean Hampton (1993, p. 153) writes, "If Aristotle is right, the harmful preferences of people not only toward themselves but also toward others cannot be considered authentic preferences of those selves, because they are the product of people in turmoil, who cannot author preferences satisfactorily." Self-authorship, in Hampton's terms (p. 155) involves a decision to develop interests and projects that are "ones *you* want, and not ones that others prefer that you want (and perhaps try to persuade you to want)." Hampton argues that the latter sort of interests can be other-regarding in an objectionable way.

[18] The account of servility presented in Hill (1991) appears to understand servility as a stable, broad-based character trait. However, we can work with a concept of servility that characterizes a type of action in the manner suggested above in much the same way that the concept of courage can individuate a type of action as well as a type of character. We cannot conclude on the basis of the empirical evidence surveyed here that the individuals in (§1) were servile as a matter of character; nevertheless we can correctly describe their behavior as servile.

Borrowing from Thomas Hill (1991), let us say that servility involves a deferential attitude towards others produced by ignorance, misunderstanding, or devaluation of one's status in relation to others.[19] Hill speaks in terms of a devaluation of one's "moral rights;" however, an ecumenical notion of servility can be articulated in terms of a more general idea of equality. Most modern moral theorists accept some conception of the basic moral equality of all persons, although particular specifications of this idea vary across theories. Servility is not a matter of underestimating one's talents or accomplishments; rather servile individuals consign themselves to a subordinate status.

Hill suggests that servility is the mirror image of selfishness. (1991, p. 12)   When acting in a servile manner, an individual arbitrarily denigrates herself vis-à-vis others; when acting in a selfish manner, an individual arbitrarily denigrates others vis-à-vis herself. Both servility and selfishness involve a failure to appreciate the basic moral equality of all persons—the conviction that *everyone* is entitled to consideration and respect.

As such, servile action displays a lack of an appropriate respect for considerations of moral equality. An individual that relegates herself to a lesser moral status not only allows herself to be disrespected, but is also poorly equipped to respect others. Hill (1991, p. 14) indicates that individuals willing to tolerate inequities in their own cases are more likely to tolerate (and participate in) such abuses when they befall others.

Hill proposes that a number of reasons can underlie servility. One might simply fail to understand that one can demand a basic level of respect from others. Or one might recognize, but undervalue, one's moral standing.

The reasons might be more prosaic. The individuals discussed in (§1) act immorally partly because they're trying to fit in. They seek acceptance, or they might just want their time to pass without conflict. They silence themselves to make things smoother and more comfortable.

---

[19] Kant (2002, p. 187) writes that servility involves "belittling one's own moral worth merely as a means to acquiring the favor of another."

Deviating from social norms is psychologically taxing. Standing up for oneself often means standing out. And the prospect of standing out can be awkward, alienating, and even terrifying.

There seems to be a similarity between the motives underlying servility and the motives many regard as essential to *moral* behavior. According to Hume (2005, p. 148), the "affection of humanity . . . alone can be the foundation of all morals." Most of us have sympathy and concern for the interests and feelings of others and this affection is part of what characterizes morally and psychologically healthy human beings.[20]

Hume says that we ought to be moral because immorality diminishes our standing in the eyes of others. But this is often precisely the problem facing an individual who acts immorally from servile motives—she cares *too much* about her standing in the eyes of others. For this reason, Hume's answer to the moral question is critically incomplete. Our need for belonging and affiliation—in brief, our sociality—is precisely what renders us more susceptible to immorality than commonly thought.[21] Milgram, for example, writes that "It is a curious thing that a measure of compassion on the part of the subject—an unwillingness to 'hurt' the experimenter's feelings—is part of those binding forces inhibiting his disobedience" (1973, p.76). Fellow-feeling is not always an ally of morality [22]

Hobbes's answer to the moral question does not apply here either. Hobbesians claim that accepting constraints on the direct pursuit of self-interest is rational because doing so facilitates advantageous cooperation. But Hobbes's moral psychology is insufficiently sensitive to our other-regarding sentiments. The individuals in (§1) are not motivated by self-interest, at least not as

---

[20] Indeed, one of the striking features of Milgram's and Zimbardo's experiments is that their subjects were, by all accounts, psychologically normal.
[21] I am grateful to an anonymous reviewer for bringing out this idea.
[22] In the limit, we have Rousseau's (1984, p. 136) description of "social man" who "lives always outside himself; he knows how to live only in the opinion of others, it is, so to speak, from their judgement alone that he derives the sense of his own existence . . . Everything is reduced to appearances, everything comes to be false and warped, honour, friendship, virtue . . . [W]e have only facades, deceptive and frivolous."

traditionally understood by Hobbes and others. None of them are driven to defect from a cooperative venture by the expectation of personal gain. They are not driven to break off cooperation at all. If anything, they are *too* cooperative. Sometimes they are motivated by a desire to *solidify* cooperation.

Once we rethink the reasons people have to be immoral, we need to rethink the reasons they have to be moral. Reasons formulated in terms of the immoralist's self-interest will not always suffice. To fully answer the moral question, we must offer individuals a reason to abstain from servility. Answers that appeal to our sense of fellow-feeling do not appreciate the ways in which fellow-feeling itself can motivate immorality; answers that appeal to enlightened self-interest assume that we are antecedently committed to pursuing our self-interest, a commitment which is absent in cases of servile immorality.

The answer I present here makes the self-destructive psychological consequences of servility salient. Self-interest presupposes a self whose interests one seeks to satisfy. One has reason to abstain from servility because one has reason to preserve a self whose interests one deems *worth* satisfying. In Korsgaard's terms, individuals have reason to abstain from servility because they have reason to preserve their practical identity, that is, "a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking" (2005, p. 101).

Unlike Korsgaard, however, I will argue that servility imperils the integrity of the *contingent* dimensions of one's practical identity: (i.e.) those particular desires, beliefs, and projects that constitute agents' conceptions of themselves. Although Korsgaard recognizes that some aspects of our practical identity are contingent, her theory of moral obligation does not account for their importance to us. Richard Joyce (2006, p. 198) argues that even if being a rational agent is necessary for one's contingent identities, it does not follow that being a rational agent is one's most valuable

12

identity. Joyce notes that his identity as a father is dependent on his identity as a man, but this does not imply that the values associated with being a man are more important than the values associated with being a father. Our contingent identities may be those we value most.

Our contingent practical identities are constituted in part by the ends we set for ourselves. To sincerely identify oneself as, say, an environmentalist, requires more than simply professing belief in a cause. To adopt this role as part of one's identity is to adopt goals such as participating in activism or reforming one's dietary habits.[23] Part of what it is to be an environmentalist is to dispose oneself to pursue certain sorts of ends. Thus, if one values one's identity as an environmentalist, one has reason to sustain one's commitment to the relevant family of ends. A concern for the integrity of the particular ends and projects that, in large part, define who we are as individuals is precisely what underwrites the following answer to the moral question.[24]

§3

Servility manifests in compliance. We can use the term 'compliant' in a technical sense. Compliance occurs "when an individual accepts influence from another person or a group in the hope of achieving a favorable reaction, or avoiding an unfavorable reaction, from the other" (Kelman and Hamilton 1989, p. 104). Milgram's subjects did as the experimenter ordered partly to elicit a favorable response from him. The Duplicitous Newcomer, unsure of her place in her community, cannot bring herself to defy the expectations of her new neighbors. Soldiers in Battalion 101 did not want to lose face with their peers.

---

[23] For further discussion of the relationship between an individual's character and her projects, see Bernard Williams (1981b).

[24] Jean Hampton (1993, p. 150) emphasizes that self-authorship involves a conception of one's "personal needs," i.e., that which "one requires as a particular personality or self, and is subjectively defined, arising from a person's decision to be a certain way, to have certain aspirations, and to undertake certain projects— all of which are up to her to determine." Hampton underscores the importance of those traits and goals that differentiate us as individuals, in contrast to traditional Kantian accounts of obligation that focus on that which we share as rational beings.

Compliant individuals adopt the socially induced behavior to produce some social effect, not because they accept its content (Kelman and Hamilton 1989, p.104). Servility therefore involves hypocrisy—acting contrary to one's attitudes. The Duplicitous Newcomer affirms Whig values but voices Tory ones. As noted, a guard in the Stanford Prison reported that comporting with his social role led him to act with a cruelty he would otherwise find unacceptable. He felt compelled to "put on a face" (49). Many of Milgram's subjects disagreed with the orders, but obeyed them anyway. They wanted to please the authority or to avoid appearing "arrogant, untoward, or rude" (Milgram 1973, p. 76).

Over time, compliance can breed internalization.[25] Unlike mere compliance, internalization involves a change in a person's private beliefs. Internalization occurs when one's sincere values conform to one's socially induced behavior.[26] After repeated acts of cruelty, many of Milgram's subjects began to manufacture reasons for their behavior. Milgram writes, "Many subjects harshly devalue the victim *as a consequence* of acting against him. Such comments as, 'He was so stupid and stubborn he deserved to get shocked,' were common. Once having acted against the victim, these subjects found it necessary to view him as an unworthy individual, whose punishment was made inevitable by his own deficiencies of intellect and character" (1975, p. 10). The phenomenon of attitudes conforming to action is well-documented. David Myers writes, "If social psychologists have proven anything during the last thirty years they have proven that the actions we elect leave a residue inside us. Every time we act, we amplify the underlying tendency . . . Our traits and attitudes *follow* our behavior" (1993, p. 123).

---

[25] On internalization and compliance, see Kelman and Hamilton (1989, p. 110).
[26] Kelman and Hamilton write, "Behavior adopted through internalization is in some way—rational or otherwise—integrated with a person's existing values. It becomes part of a personal structure, as distinguished from a set of social-role expectations. Such behavior gradually becomes independent of the external source" (1989, p.109).

Cognitive dissonance explains why internalization follows compliance.[27] Cognitive dissonance occurs when a person holds two cognitions (e.g., beliefs, desires, attitudes) that are inconsistent (cognitions X and Y are inconsistent if X implies ~Y).[28] Individuals are motivated to reduce cognitive dissonance because the psychological tension is unpleasant.[29] To continue with our example, Milgram's subject notices that the learner has done nothing to warrant mistreatment. Yet he mistreats the learner. To delete the inconsistency, the subject fabricates a reason for the mistreatment—the learner must deserve it after all. Elliot Aronson writes, "Your cognition 'I am a decent moral person' is dissonant with your cognition 'I have just committed an immoral act' . . . One way to reduce dissonance is to minimize the negative aspects of the action you have chosen (and to maximize the positive aspects)" (1972, p. 107). Milgram's subjects did exactly that: they minimized the negativity of the shocks by convincing themselves that the learner deserved them.

Zimbardo attributes the psychological transformation of his prison guards to dissonance reduction. He asks, "[P]recisely where does one's 'identity' end and one's 'role' begin? When the private self and the public role behavior clash, what direction will attempts to impose consistency take?" (1973, p. 48). As indicated earlier, the influence of the private self diminished and that of the public role grew. "It was remarkable how readily we all slipped into our roles [Zimbardo himself was the mock prison superintendent], temporarily gave up our identities and allowed these assigned roles and the social forces in the situation to guide, shape, and eventually to control our freedom of thought and action" (Zimbardo 1973, p. 46).

For another example of how public behavior can restructure an individual's private beliefs, consider the phenomenon of 'character invasion.' In theatre, character invasion occurs when an actor's stage character affects his personality off-stage. Acting, like servility, involves dissonance

---

[27] Bicchieri (2006, p. 194) also suggests a link between internalization and cognitive dissonance.

[28] Cognitive dissonance can occur in a variety of ways. However, for simplicity's sake, I leave some aside.

[29] See Aronson (1972, p. 93).

between one's private beliefs and one's public behavior. In an article on character invasion, an actor discusses his experience with the phenomenon:

> One of our own interviewees, Tom, described a moment in an undergraduate rehearsal in which, following an intense exercise, "I could not find any way to maintain a hold on who I was. . . . There was no sense of an 'I' as in 'me, Tom.' It was all a sense of an 'I' as in 'me, [the character].'" Furthermore, Tom found his character's manipulative behavior carried over into his real life. Additionally, before the first show closed, he went immediately into rehearsal for a second show. Consequently, Tom said, "I was still having a problem finding who I was, so there was no 'me' to . . . build the other character from. So I was like three characters all fighting for the same body (Burgoyne et al. 1999, p. 164).

Actors can become like the characters they portray. Invasion can be so powerful that avoiding it is a *moral* issue in theatre pedagogy (Burgoyne et al. 1999).

Similar effects can be produced without even acting against one's private beliefs—experimental evidence confirms that inducing individuals to simply speak against their beliefs reduces their confidence in those beliefs.[30] We should be unsurprised to learn that after consistently expressing claims contrary to her own beliefs, the Duplicitous Newcomer undergoes a real change in her beliefs. If she repeatedly defends a view, it is natural that she will, over time, become more confident in its assertion. For instance, if she constantly defends the Tory leader, states the reasons why he ought to be supported, and argues on his behalf, her opposition will eventually fade and perhaps even transform into support.[31]

Even if one does not experience the kind of sustained pressure that can produce a transformation of one's practical identity, servility may still produce *apathy* towards one's practical identity. Milgram's obedient subjects disregarded the learner's desire to end the experiment, allowing it to be overridden by the demands of the authority. Acting against the learner's desire to be released led the subject to devalue the learner's desire to be released. The thought "I am disregarding his wants as unimportant and unworthy of consideration" is dissonant with the thought

---

[30] See, for instance, Mussweiler, et al. (2000).

[31] I do not mean to suggest that individuals ought to accept the validity of their beliefs uncritically, ignoring countervailing evidence. But the Duplicitous Newcomer's purpose in speaking against her political beliefs is not to question their validity; rather her purpose is to signal to others that she belongs, i.e., that she is not questioning *their* beliefs.

"His wants are important and worthy of consideration." To reduce this dissonance, subjects minimized the importance of the learner's protests.

Yet the subject also disregarded his *own* desire to end the experiment, allowing it to be overridden by the demands of the authority. To reduce the dissonance between his disregard of his wants and his belief that his wants are important, one might expect the subject to minimize the importance of his wants, at least relative to the demands of the authority.[32] Acting against the learner's desires led the obedient subject to judge the learner's desires as unworthy of consideration. Acting against one's own desires might lead one to judge one's own desires as unworthy of consideration.

Similarly, when the Duplicitous Newcomer first lies about her politics, she still has a commitment to her views—there's an internal protest. But over time, duplicity can weaken her resolve. Lies and compromises wear down the internal protests. She becomes apathetic to the point that she just doesn't care about politics anymore. She no longer sees her political ideals as ends worth advancing. This kind of slide into indifference can also be explained by cognitive dissonance. It is called the 'sour grapes' syndrome: the interests that a choice does *not* satisfy become proportionately less important to the chooser.[33]

§4

My aim in this section is to argue that individuals normally have reason to abstain from servility. The argument has two premises. The first is the proposition:

> (1) If φing frustrates subject S's ends, S has (defeasible) reason to abstain from φing.

---

[32] There is evidence to suggest that we diminish the importance of values that are precluded by our behavior. "If [a person] has tried to quit smoking and has failed, he is committed to smoke. Thus, he becomes less intensive in his belief that smoking is dangerous" (Aronson 1972, p. 95). If you are unable to stop smoking, smoking seems less harmful. Perhaps if you are unable to assert yourself, then the failure to do so seems less harmful.

[33] On the sour grapes syndrome, see Elster (1993, p. 54).

I simply assume that most will accept premise (1), although not all will. Note that the premise does not state that an agent has a reason to abstain from φing if she *believes* that φing frustrates her ends. Here I follow Bernard Williams's "sub-Humean model" of practical reasoning (Williams 1981a, p. 102). For Williams, whether or not an agent has a reason for action depends on facts, not merely beliefs, about whether the action will achieve or frustrate the agent's end.

Williams considers a thirsty agent who believes that a glass of liquid contains gin, when the glass actually contains petrol. Williams claims that the agent has reason not to drink the liquid because doing so does not in fact serve his end, even though he believes it does. Agents can have reasons for action that are contrary to their *perceived* reasons for action.

Here is the second premise of the argument:

(2) Servility frustrates S's ends.

In the previous section, I argued that servility causes undesired self-alteration. Servility requires acting in contradiction to our beliefs and evaluative attitudes. Evidence from social psychology suggests that agents reduce this cognitive dissonance by unconsciously adapting their attitudes to their actions, or by extinguishing such attitudes. The self-alteration produced by dissonance reduction frustrates our ends (that is, our ends prior to the dissonance reduction) by altering our evaluative attitudes, thereby altering what we intend to do. Agents have a reason to abstain from behavior productive of this kind of self-alteration in the interest of preserving the relevant intentions. This idea can be illustrated by two examples.

Suppose at lunch the agent has the end of eating a steak at dinner. Unknown to the agent, the hamburger on the lunch menu will extinguish his hunger for the steak. Eating the hamburger at lunch will frustrate his end of eating the steak at dinner. Thus, the agent has a reason to abstain from eating the hamburger at lunch, even though he may not recognize this reason.

Servility presents a parallel case. Suppose the Duplicitous Newcomer moves to her new home in January. At this time, she has the end of voting for the candidate that she favors upon undergoing an impartial, well-informed deliberation, which turns out to be the Whig candidate. The conclusion from the previous section states that repeatedly lying about politics to ingratiate herself can alter the Duplicitous Newcomer's political views. Let us stipulate that this change includes a reversal of opinion about the candidate she chooses to vote for. Iterating a pattern of servile behavior beginning in January would frustrate the Duplicitous Newcomer's end of voting for the Whig candidate in November. Thus, the Duplicitous Newcomer has a reason to abstain from that behavior. Premises (1) and (2) entail the conclusion: (3) S has (defeasible) reason to abstain from servility.[34]

Although many accept the instrumentalist claim that ends are reasons for the actions that serve them, not all do. Perhaps we can sketch an argument to abstain from servility that speaks to those who reject the Humean account of practical reason. Consider that the Duplicitous Newcomer abandons her end of voting for the Whig candidate not because she has undertaken a critical survey of the relevant evidence, but because she has been transformed by the non-rational resolution of cognitive dissonance. She has no reason to regard her later decision to vote for the Tory candidate as *justified*. Thus, a case might be made against servility on the grounds that individuals ought not to revise their ends in the absence of justification for such a revision.[35]

This sketch allows that compliance as such need not be objectionable or servile. Indeed, compliance can be a vital and beneficial part of human social life. What is objectionable about many cases of compliance is individuals' willingness to comply in the absence of any compelling reason

---

[34] Of course, because this argument relies on empirical assumptions about human psychology, it should not be construed as asserting that agents *necessarily* have reason to abstain from servility.
[35] I am grateful to an anonymous reviewer for suggesting this argument.

and often in spite of compelling reasons *not* to comply.[36] But when individuals comply, for instance, as a means to achieving some independently valuable end like learning from an expert, their action need not be servile. Thus it is not the case that we should never comply, but rather that when we do comply, we should do so for the sake of some further independently valuable end.[37]

§5

Perhaps I've overstated my case. We may have reason to abstain from iterating a pattern of servile behavior, but what's the harm in occasionally telling lies to get along easier? Or belittling oneself to curry someone's favor once or twice? Aren't cases of extreme self-alteration rare? Our identities are more stable than my discussion implies.

On their own, small concessions to social pressure probably will not do much damage. The problem, however, is that each concession makes subsequent concessions more likely. The brutality of the individuals in Milgram's and Zimbardo's experiments and at Abu Ghraib was made possible in part by the gradual escalation of the severity of the acts.[38]

Consider that Milgram's subjects began by applying small shocks of fifteen volts. The learner did not protest these. The ensuing shocks increased by tiny increments. The subjects complied five times before the learner first grunted (1975, p. 23). The subjects complied ten times before the

---

[36] The psychology of conformity is more complicated than what is indicated by my examples alone. An anonymous reviewer notes, for example, that some individuals hid their Jewish identity during the Inquisition for years without losing it and that, more generally, people are sometimes strengthened in their resistance to a public culture in virtue of being forced to publicly conform. However, the individuals in these cases differ from those in the cases I present in virtue of conforming for the sake of some further valued end. My examples feature individuals who seek conformity for its own sake. To be sure, this need not be the case for all examples of conformity. Indeed, conforming for instrumental reasons does not seem objectionable in itself; conformity is disconcerting primarily when individuals defer to others for no further purpose.

[37] A similar point can be made about the sort of person from whom we seek approval. Milgram (1973, p. 76) notes that his subjects enjoyed the feeling of pleasing the experimenter. But the experimenter had done little to demonstrate that he was the sort of person the subjects should be interested in pleasing. Milgram even replicated the experiment at an office building in Bridgeport, Connecticut without any visible affiliation with Yale in order to discover the role Yale's prestige played in eliciting obedience. He found that "the level of obedience in Bridgeport, although somewhat reduced, was not significantly lower than that obtained at Yale" (1975, p.69). There is an important distinction between an indiscriminate desire for approval and a reasoned desire for the approval of those one has reason to respect.

[38] For a useful discussion of the so-called "foot in the door" technique, see J. Freedman and S. Fraser (1966).

learner protested (1975, p. 23). Compliance with innocuous orders made future compliance with harsher orders—and their concomitant internalization—easier. Had Milgram's subjects' first shock caused the victim to scream in protest, the subjects would probably have stopped right there. But compliance to orders demanding extremely severe shocks was made possible by previous compliance to orders demanding lesser (but increasingly severe) shocks.

Similarly, the Stanford Prison Experiment illustrates how minor acts of compliance can snowball. Before the experiment, Guard A wrote in his diary, "As I am a pacifist and nonaggressive individual I cannot see a time when I might guard and/or maltreat other living things" (54). Indeed, after the orientation meeting he added, "I doubt whether many of us share the expectations of 'seriousness' that the experimenters seem to have" (54). But within five days of playing the role of guard, the subject was locking Prisoner 416 in solitary confinement and trying to force feed him simply because he refused to eat dinner.

One guesses that on the first day, Guard A would not have even contemplated force feeding dissidents and locking them in solitary confinement. But after repeated acts of compliance to the basic and fairly innocuous norms of guard behavior (e.g. not smiling at prisoners, refusing their requests for cigarettes, contradicting them, etc.), compliance to more egregious norms became more natural.

Our identities are stable, but not invulnerable. Milgram's obedient subjects, for instance, probably found that their hostilities faded quickly. Their change was shallow. But aggregating minor acts can cause dramatic changes over time. In less than one hour, Milgram's subjects went from delivering shocks that (they believed) barely tickled the learner to delivering shocks that left him screaming in pain, and ultimately silent. In less than one week in Zimbardo's experiments, pacifists were brutalizing mock prisoners.

These considerations speak in favor of a resolute commitment to abstain from servility. Korsgaard writes, "[Y]ou can stop being yourself for a bit and still get back home, and in cases where a small violation combines with a large temptation, this has a destabilizing effect on the obligation . . . [W]e must commit ourselves to a kind of second-order integrity, a commitment to not letting these problems get out of hand. We cannot make an exception 'just this once' every time, or we will lose our identities after all" (2005, p. 103). We should not underestimate the cumulative effects of servility over the months, years, or a lifetime.

There is an Aristotelian point here: we are what we repeatedly do. When we tell a friend a painful truth, we become a little more honest. When we give into fear, we become a little more cowardly. When we allow ourselves to be manipulated, we become a little more servile. Repeated actions become habits and habits become the cells of character.

§6

I have argued that in order to justify morality from the first-person perspective, we must offer individuals reasons to abstain from servility that appeal to their own motivational structures. Answers to the moral question that appeal to our fellow-feeling or our self-interest do not adequately address the fact that we sometimes act wrongly out of fellow-feeling and to the acknowledged detriment of our personal interests. My argument holds that we should refrain from servile behavior on the grounds that it can corrode our contingent practical identity, an outcome we have reason to avoid.

This paper is an attempt to redress a deficit in prevailing treatments of the rationality of morality. Servility's presence in this literature is disproportionate to its presence in our moral lives. To be sure, people act immorally out of perceived self-interest or because they fail to respect others. But they also act immorally because they fail to respect themselves. To fully substantiate the

rationality of morality, we must address those individuals who transgress out of a failure to affirm and appreciate their own worth.

References

Aristotle, *Nicomachean Ethics*. Translated by Roger Crisp (Cambridge: Cambridge University Press, 2005).

E. Aronson, *The Social Animal* (San Francisco: W.H. Freeman, 1972).

C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (New York: Cambridge University Press, 2006).

S. Burgoyne, K. Poulin, and A. Rearden, "The Impact of Acting on Student Actors: Boundary Blurring, Growth, and Emotional Distress," *Theatre Topics* 9, no. 2 (1999): 157-179.

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

D. Copp, "Moral Naturalism and Three Grades of Normativity," in *Normativity and Naturalism*, ed. P. Schaber. (Frankfurt: Ontos-Verlag, 2004), 7-45.

R. Crutchfield, "Conformity and Character," *American Psychologist* 10 (1955): 191-198.

J. Elster, *Political Psychology* (Cambridge, Cambridge University Press, 1993).

S. Fiske, L. Harris, and A. Cuddy, "Why Ordinary People Torture Enemy Prisoners," *Science* 306 (2004): 1482-1483.

J. Freedman and S. Fraser, "Compliance Without Pressure: The Foot-in-the-Door Technique," *Journal of Personality and Social Personality* 4 (1966): 195-202.

B. Gert, *Morality: Its Nature and Justification* (New York: Oxford University Press, 1998).

A. Gewirth, *Reason and Morality* (Chicago: University of Chicago Press, 1981).

J. Hampton, "Selflessness and the Loss of Self," *Social Philosophy and Policy* 10 (1993): 135-165.

T. Hill, "Servility and Self-Respect" in his *Autonomy and Self-Respect* (New York: Cambridge University Press, 1991), 4-18.

D. Hume, *An Enquiry Concerning the Principles of Morals*. (Oxford: Oxford University Press, 2005).

R. Joyce, *The Evolution of Morality* (Cambridge, MA: The MIT Press, 2006).

I. Kant, *The Metaphysics of Morals*. Translated by M. Gregory (Cambridge: Cambridge University Press, 2002).

H. Kelman, *Crimes of Obedience: Toward a Social Psychology of Authority and Responsibility* (New Haven: Yale University Press, 1989).

C. Korsgaard, *The Sources of Normativity* (New York: Cambridge University Press, 2005).

S. Milgram, "Behavioral Study of Obedience," *Journal of Abnormal and Social Psychology* 67, no.4 (1963): 371-378.

S. Milgram, "The Perils of Obedience," *Harper's* 247, no. 1483 (1973).

S. Milgram, *Obedience to Authority* (New York: Harper Perennial, 1975).

T. Mussweiler, F. Strack, and T. Pfeiffer, "Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility," *Personality and Social Psychology Bulletin* 26 (2000): 1142-1150.

D. Myers, *The Pursuit of Happiness* (New York: Avon Books, 1993).

R. Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

H. Pritchard, "Does Moral Philosophy Rest on a Mistake?" *Mind* 21 (1912): 21-37.

J. Rousseau, *A Discourse on Inequality.* Translated by Maurice Cranston (London: Penguin Classics, 1984).

D. Schmidtz, "Self-Interest: What's in It for Me?" in *Self-Interest*, eds. E. Paul, F. Miller, and J. Paul (Cambridge: Cambridge University Press, 1997), 107-121.

R. Shafer-Landau, *Moral Realism: A Defence* (Oxford: Oxford University Press, 2003).

C. Vogler, *Reasonably Vicious* (Cambridge, MA: Harvard University Press, 2002).

B. Williams, "Internal and External Reasons," in *Moral Luck* (Cambridge: Cambridge University Press, 1981a), 101-113.

B. Williams, "Persons, Character, and Morality," in *Moral Luck* (Cambridge: Cambridge University Press, 1981b), 1-19.

P. Zimbardo, *The Lucifer Effect* (New York: Random House, 2007).

P. Zimbardo, et al. "The Mind is a Formidable Jailer: a Pirandellian Prison," *The New York Times Magazine* (April 8, 1973): 38-60.