
Problems in the Theory of Ideology

Joseph Heath

One of the most persistent legacies of Karl Marx and the Young Hegelians has been the centrality of the concept of "ideology" in contemporary social criticism. The concept was introduced in order to account for a very specific phenomenon, namely, the fact that individuals often participate in maintaining and reproducing institutions under which they are oppressed or exploited. In the extreme, these individuals may even actively resist the efforts of anyone who tries to change these institutions on their behalf. Clearly, some explanation needs to be given of how individuals could systematically fail to see where their interests lie, or how they might fail to pursue these interests once they have been laid bare. This need is often felt with some urgency, since failure to provide such an explanation usually counts as *prima facie* evidence against the claim that these individuals are genuinely oppressed or exploited in the first place.

There is of course no question that this kind of phenomenon requires a special explanation. Unfortunately, Feuerbach, Marx, and their followers took the fateful turn of attempting to explain these "ideological" effects as a consequence of *irrationality* on the part of those under their sway.¹ While there are no doubt instances where practices are reproduced without good reason, the ascription of irrationality to agents is an explanatory device whose use carries with it considerable costs, both theoretical and practical. As a result, it should be used only in the last resort. In this paper, I will argue that many of the phenomena traditionally grouped together under the

category of “ideological effects” can be explained without relinquishing the rationality postulate. I will try to show that agents *can* rationally engage in patterns of action that are ultimately contrary to their interests, and that they can rationally resist changing these patterns even when the deleterious or self-defeating character of their actions has been pointed out to them.

I think that an approach such as this, one that is sparing in its ascription of irrationality and error, has two principal advantages. First, it allows one to engage in social criticism while minimizing the tendency to insult the intelligence of the people one is trying to help—whether it be the working class, women, or the subaltern. This may reduce the tendency exhibited by some members of these groups to reassert their autonomy precisely by rejecting the critical theory that impugns their rationality. The second major advantage is also practical. The vast majority of oppressive practices are not reproduced because people have false or irrational beliefs. As a result, simply persuading people to change their beliefs will have no tendency to change the underlying mechanism through which the practices are reproduced. Even worse, the institutions in question may acquire the appearance of being impervious to social criticism, simply because people continually criticize them, and yet nothing changes. Correctly diagnosing the mechanism through which these institutions are reproduced has the potential to suggest new strategies for social change.

1 Ideology and Irrationality

One might begin by asking what the big problem is with the assumption that people are behaving irrationally. After all, everyone knows that people make mistakes, and do things without thinking. If people are acting in a way that is making their own lives miserable, it seems likely that they are making some kind of mistake. When the peasant rallies to the defense of his feudal lord, or the hostage begins to promote the goals of her kidnappers, we are likely to think that these people are behaving in a muddle-headed way. We may even come up with a name for what has muddled them up, calling it “Christianity,” or “Helsinki syndrome.” When they persist in this behavior, even after having been told that they are suffering from

one of these ailments, we might start to think that the problem is even deeper, and that something has impaired their ability to assess the information they have been given. We begin to think that they are not just mistaken, but in the grip of some deeper form of irrationality. This diagnosis seems fairly intuitive—so what is the big problem?

The problem is the one raised by Donald Davidson in his famous critique of “conceptual schemes.”² Davidson’s argument is roughly as follows: there is no fact of the matter about what people mean by what they say. The meaning of their utterances is determined by the best interpretation that hearers confer upon them. However, the meaning that I ascribe to a person’s utterances depends in a crucial way upon the set of beliefs that I take that person to hold true. For instance, when people talk about meeting on “Thursday,” I can only figure out which day they are referring to by assuming that they share with me the belief that today is Tuesday. If I thought they believed that today was Monday, I would start to think that they meant Wednesday when they said “Thursday.” But since we can only find out what peoples’ beliefs are by asking them, and since they can only express their beliefs by putting them in the form of sentences that in turn require interpretation, any particular interpretation that we might confer upon a person’s utterances is massively underdetermined by the evidence available to us. There are an infinite number of ways to interpret anyone’s speech, each supported by the ascription of a different set of beliefs.

But then how do we ever understand one another? Davidson argues that all interpretations are constrained by a principle of charity. The best interpretation is the one that ascribes the most reasonable set of beliefs to that person, which is to say, the one that maximizes the number of true beliefs the person is thought to hold. From the standpoint of the hearer, this means that the best interpretation is the one that is consistent with the highest level of agreement between the speaker and the hearer. This requirement of charity is not a methodological assumption, it is a constitutive principle. To interpret someone is to interpret that person charitably—if you are not interpreting them charitably, then what you are doing simply does not count as interpretation.

Davidson's argument has the effect of eliminating a certain Cartesian form of skepticism. Imagine a person transported from the real world, in which she has true beliefs about the objects around her, into a dream world, in which she is fooled into retaining the same beliefs about her environment, despite the fact that the corporeal objects have all disappeared. Her beliefs, which were once predominantly true, are now predominantly false. But how could she ever know? This is the basic structure of the Cartesian thought-experiment. Davidson's argument exposes a crucial presupposition. It is simply assumed that the person's beliefs are *the same* in both the real world and the dream world, and in particular, that they have the same content. However, since beliefs are propositional attitudes, that is, interpreted sentences of a natural language, this amounts to the assumption that the meaning of linguistic expressions is determined by something that remains invariant in the transition from the real world to the dream world. But according to Davidson, the meaning of expressions is determined by the practices of interpretation of a linguistic community. As a result, the meaning of the person's beliefs would change in the transition from the real world to the dream world, in such a way that they remained predominantly true. Why is this? Interpretation governed by the principle of charity requires as much, for it requires a practice of interpretation that treats the speaker as having predominantly true beliefs. If a particular interpretation made someone come out as having too many false beliefs, all this would do is supply evidence against that interpretation.

To take a real-life example of this principle in action, consider the following episode from the history of ethnography. Lucien Lévy-Bruhl famously suggested that he had discovered the existence of "prelogical" cultures.³ He found that his subjects persistently made contradictory statements, incoherent observations, and generally believed ridiculous things. Later generations of ethnographers, of course, returned to these societies somewhat skeptical about this claim, and quickly discovered other ways of interpreting the kind of utterances that had stumped Lévy-Bruhl, interpretations that made the "natives" come out sounding a lot more reasonable. For example, by distinguishing between expressions that were meant literally and

those that were meant metaphorically or figuratively, a substantial portion of the "contradictions" Lévy-Bruhl uncovered could be dismissed. The Davidsonian point is this: these later interpretations were better than Lévy-Bruhl's, not because they came closer to what the people "really" meant, but because they were more restrained in their ascription of error. They were right precisely *because* they made the natives sound reasonable. What other evidence could there be for the correctness of an interpretation? There is no "third thing" against which the two interpretations can be compared—only more interpretations.

The more general problem is this: suspending the assumption that people are by-and-large reasonable, and that their beliefs are predominantly true, removes the only constraint that prevents one from interpreting their utterances as meaning anything at all. The problem then is not that one can no longer construct a plausible explanation of their behavior, but that one can construct too many plausible explanations, and it is hard to rule any of them out. This means that the critical theorist can only go so far in ascribing irrationality and error to people. Once she crosses a certain threshold, this ascription of error stops being an exposé of their mistakes, and starts to count as evidence against the proposed interpretation of their behavior. It starts to suggest that, rather than having uncovered a massive, all-encompassing ideology, she has simply failed to understand what it is that people are doing. The interpretation that appeals less heavily to ideology then wins, for that very reason—it appeals less heavily to ideology.

This is a classic problem in the history of Frankfurt School critical theory. The "grandfathers" of critical theory, Marx and Freud, both developed theories that diagnosed very widespread error in popular belief.⁴ In so doing, they ran the risk of undermining themselves. Attributing massive error to people significantly expands the range of motives and beliefs that *can* be ascribed to them. This makes it very hard to show that any one interpretation is better than another. For this reason, Marx and Freud both had to bring in some kind of an "external" theory—historical materialism, psychoanalysis—that could be used to assess the merits of their interpretations. The problem is that

this standard then needed to be grounded in something “above” the fray of competing interpretations and theories. It had to be more than just the best interpretation of some data set, since it purported to provide the criterion of interpretive adequacy.

Marx and Freud tried to handle this problem by claiming that their metainterpretive theories were “scientific,” and thus presumably justified in a foundational manner. Unfortunately, there was no particular merit in either claim; both merely traded on the prestige associated with “scientific” inquiry. As a result, critics were quick to turn the critical theories of Marx and Freud against the critics. Karl Mannheim argued that if classical economics was merely the ideology of the bourgeoisie, then historical materialism was nothing more than the ideology of the proletariat, and therefore no more true.⁵ Similarly, psychoanalysts were quick to suggest that Freud’s instinct theory, which was to serve as a guide for the construction of interpretations, was itself nothing but a manifestation of his own unconscious desires.⁶ What these criticisms have in common is that they deny any privileged status to the metainterpretive theories, thereby drawing them down to an interpretive level. But once this is done, then the Davidsonian problematic begins to loom large—it becomes very unclear what is supposed to make one critical theory more plausible than the next.

Much of the history of critical theory in the twentieth century can be seen as an attempt to work around this problem. However, one of the things that is seldom questioned in this body of work is the very basic assumption that when people act in a way that is contrary to their interests, they must somehow be acting irrationally. I would like to suggest that this is often not the case. While people do sometimes make mistakes and get confused, this is more the exception than the norm—especially when it comes to their core economic interests. I will try to show that individuals often get outcomes they don’t want, not because they have chosen wrongly, but because their actions combine with those of others in undesirable ways. Thus greater attention to the structure of social interaction reduces the need for a theory of ideology. In the next three sections, I present three different ways in which agents can rationally choose to perform actions that are, in some sense, contrary to their interests.

2 Collective Action Problems

The most common error that critical theorists have made, in my view, is to mistake the outcome of a collective action problem for an effect of ideology. Collective action problems arise in situations where agents can best pursue their own goals and projects only by imposing some kind of a cost upon others. The prisoner’s dilemma is the classic example: each suspect can reduce his own expected jail time by turning in his partner. Doing so, however, increases the amount of jail time that his partner must serve. As a result, both suspects turn each other in, and so both serve more jail time than either would have had they remained silent.

Many interactions involving large numbers of people have precisely the same structure. For example, telephone companies did not used to bill their customers for individual calls to directory assistance. Instead, customers paid for the directory assistance service as part of their basic monthly package. The problem with this arrangement is that it generates overuse of the service, since the cost of serving any individual caller is paid by *all* of the firm’s customers. So individuals who were too lazy to look up a number could get someone else to do it for them, while effectively displacing the cost of this action onto others. But when everyone does this to each other, everyone winds up using more directory service, and paying more for it, than anyone actually wants to. As a result, when phone companies switched to charging individuals directly for calls to directory assistance, the volume of calls dropped dramatically. (In a trial run in Cincinnati, imposition of a \$0.20 per call charge reduced the average number of directory assistance calls from 80,000 to 20,000 per month. As a result, average residential telephone rates dropped by \$0.65 per month.⁷)

In the case of directory assistance, people get into a collective action problem in part because the interaction is completely anonymous. The same dynamic, however, can show up in face-to-face interactions. For example, if a large group of people all know that a restaurant bill is going to be divided evenly among them, regardless of what they order, individuals will often order a more expensive meal than they would if they were eating alone. But if everyone does this,

then everyone winds up spending more than they want to. If you are eating out in a party of six, then the \$12 appetizer that you order only costs you \$2 (and you are going to pay for part of your neighbor's appetizer, regardless of whether you order one for yourself, so you might as well indulge). For the same reason, individuals in large groups tend to shortchange the house staff on the tip. Everybody undertips on their portion of the bill, because instead of making you look *very* cheap, it just makes everyone in your party look a little cheap. Of course, when everyone does it, it just makes everyone look *very* cheap. (This is a common outcome, and is in fact the reason why most restaurants impose a mandatory minimum gratuity on large parties.)

The most significant thing about these collective action problems, from the standpoint of critical theory, is that agents often have a hard time getting out of them, even if they realize that they are engaging in collectively self-defeating behavior. The reason is that the mere recognition that the outcome is suboptimal does not change the incentives that each individual has to act in a way that contributes to it. Even if I realize that I shouldn't overuse directory assistance, it doesn't mean that my phone bill will get any lower if I stop. It is only if *everyone* stops that I will begin to see a difference. But I have no control over what everyone else does (and furthermore, if everyone else stops overusing the service, and I continue to do so, then I am even better off). Not only that, but even individuals who are not actively seeking to displace costs onto others may have to do so just to avoid being exploited by those who are. When I sit down with my friends in a restaurant, and they start out by ordering a round of expensive drinks, then I might as well have one too. That way instead of just paying for their drinks, at least I get to have one too.

The point is that collective action problems can be extremely hard to shake. (That's why they're called collective action problems—in order to change the interaction pattern, you need to get everyone to stop doing what they have been doing.) As a result, one good clue that people are stuck in one of these equilibria is that everyone knows what the problem is, but nothing ever changes. For example, it has been common these past years to hear complaints about the way a "media circus" develops around certain events or stories, such as the

Lewinsky-Clinton scandal, or the O.J. Simpson trial. One of the most commonly criticized characteristics of this pattern is the way that coverage achieves a "saturation" level—the clearest instance being when every major network is showing exactly the same thing. This is clearly a suboptimal outcome—if one channel is providing twenty-four-hour live coverage of a particular story, then there is no point in having the others do the same. It would be better if they showed something else, so that viewers not interested in that story would have something to watch. The same thing applies when every news program covers exactly the same five or six stories in their evening broadcast, to the neglect of other newsworthy subjects.

In any case, what is interesting about this criticism is that it is not just circulated in the broader public sphere. When the journalists who are actually providing the "saturation" coverage are asked for their views, they also often say that the situation is ridiculous, that there are interesting stories being ignored, and so forth. In other words, the people who are covering the news stories often think that they should be covering something else. Thus the problem is not that the members of "the media" have mistaken priorities, or a poor understanding of what should be on television. They can see perfectly well what is going on. The problem is that they are stuck in a suboptimal equilibrium. Stations compete with one another for viewers. Imagine a two-channel universe in which a big news story breaks. Suppose that 70 percent of television viewers will be interested in that story. If a station knows that its rival down the street will be covering it, then it has a choice of covering something else, and attracting 30 percent of the viewers, or providing exactly the same coverage, and attracting 35 percent of viewers. Both stations reason the same way, and so both provide exactly the same coverage. The result is simply a waste of one broadcast frequency.

As a result, if everyone knows what the problem is, but nothing ever changes, there is good reason to suspect the existence of a collective action problem. Similarly, if people know that a certain social change would be *in their interest* (broadly construed), this does not mean that they will have an incentive to do anything about it. I may know that it is in *our* interest, as telephone rate-payers, to use directory assistance in moderation, but that does not make it in *my* interest

to do so. In the same way, journalists may recognize that their entire profession loses credibility when they pursue sensational or lurid stories—and so it is not in *their* interest—but it may still be in the interest of each individual reporter, each individual news organization, to do so (since it is possible to increase your *share* of viewers even when the total *number* drops—exactly the same logic underlies “negative” campaign ads).

From the outside, then, it may look as if people are simply confused about where their interests lie, that they are in the grip of some ideology. But upon closer examination, they turn out to be perfectly rational. They may even join the critical theorist in lamenting the sad consequences of their own actions.

This analysis invites us to look back on some of the classic cases of ideology to see if something similar might not be going on there as well. Take the working class, for instance. Once it was decided that the workers would be better off under communism than under capitalism, many theorists simply assumed that workers would go out and overthrow the system. The fact that they failed to show up at the barricades was felt to require some explanation. Ideology was the most popular candidate. So Marx suggested that they were the victims of commodity fetishism—they mistook the social relations between individuals for objective relations between things, and so became convinced that the existing economic order was immutable. However, after half a century of Marxist critique, the working class still failed to make a revolution. Theorists began to suspect that a deeper, more insidious form of ideology was at work. The most popular diagnosis was consumerism—workers had become seduced by the materialistic values of late capitalism, and so failed to support the revolution because of a mistaken belief that they enjoyed living in suburban houses, using labor-saving appliances, eating TV dinners, and so forth.

These social critics simply failed to see the more obvious explanation. Revolutions are risky business. Setting up picket lines, not to mention barricades, is tiresome, difficult, often cold, and sometimes dangerous. Even if it were in the interests of the working class to bring about a socialist revolution, this does not make it in the interest of each individual worker to help out. There is no point going to

the barricades unless thousands of your comrades intend to join you, but if thousands of your comrades are going anyhow, no one will miss you if you stay home.⁸ Revolutionary fervor can generate the solidarity needed to overcome this collective action problem in some instances. But in general there is no reason to think that workers will show any more solidarity with one another than phone company customers. And broad segments of the working class have consistently shown themselves willing to free-ride off each others' collective achievements—this is why unions usually seek legally enforced “closed-shop” arrangements.⁹

Now consider a more controversial example. We often hear the complaint that cosmetics companies, the diet industry, plastic surgeons, and so forth, exploit women. In the mid-nineties, women in the United States spent around \$20 billion a year on cosmetics, a sum that could have been used to finance 400,000 day care centers, or 33,000 battered women's shelters, or 50 women's universities, and so on.¹⁰ This is clearly a suboptimal outcome. Furthermore, the fact that men (who earn more on average) spend only a fraction of this amount maintaining their appearance, and do not suffer much anxiety over their physical condition, adds insult to injury. The difference is also widely felt to perpetuate a set of gender roles that are disadvantageous to women: it encourages the identification of male status with money, female status with beauty; it perpetuates the idea of the female sexual role as passive, and so on. Thus feminists have for a long time argued that women need to free themselves from the dependence upon beauty, and the beauty industry.

But what has become most striking about this critique is that even though the vast majority of women accept it, it has little bearing on their personal conduct. Plenty of schoolgirls can explain to you how the cosmetics industry exploits women, but this doesn't stop them from wanting to wear lipstick. Once they graduate from university, many can tell you all about the double standards of our culture, the evils of distorted body image and the objectifying male gaze, while still counting calories and drinking skinny lattes. (Companies like Kellogg's have even used the feminist critique to sell low-fat breakfast cereal to women, under the slogan “look good on your own terms.” Their goal was to allow sophisticated female consumers to purchase

diet products without having to worry that others would think that they were merely victims of the diet industry. It allowed them to be ironic victims.)

These observations have led many feminist theorists to suggest that there must be an even more insidious form of ideology at work. If women understand the structure of their oppression, and they can see how the cosmetics and fashion industry actively exploit them, then they must be out of their minds to drop a hundred dollars on the latest moisturizer. Naomi Wolf basically suggests as much, when she describes how, "to reach the cosmetics counter, [a woman] must pass a deliberately disorienting prism of mirrors, lights and scents that submit her to the 'sensory overload' used by hypnotists and cults to encourage suggestibility."¹¹ She claims that women experience an "unconscious hallucination," that female consciousness has been "colonized," that women have been "stunned and disoriented" by changing gender roles, and so forth.¹² In short, they are not acting rationally. Why are they acting so dumb? The answer is ideology: "Women are 'so dumb' because the establishment and its watchdogs share the cosmetics industry's determination that women are and must remain 'so dumb.'"¹³

However, the very fact that everyone has heard this critique a hundred times and yet nothing ever changes, suggests that what we are dealing with is a collective action problem, not a problem of ideology. This is often overlooked in the case of beauty, because the literature has a tendency to focus on the role of ideals, or archetypes, in setting the standards of beauty. As a result, it distracts from the fact that beauty has an inherently competitive structure. Although standards of beauty vary from culture to culture, every culture has *some* kind of beauty hierarchy. People derive very significant material and social advantages from their position in this hierarchy. There is a sense in which beautiful people live in a world that is very different from the one that ugly people live in (consider how different the high school experiences can be, and how much beauty has to do with that). Apart from the fact that in a traditional patriarchal marriage market, a woman can gain significant material advantages from her beauty, there is also the fact that, at the level of face-to-face interaction, beautiful people simply get treated better than not-so-beautiful people.

People can improve both their quality of life and their material position by moving up a few levels in the beauty hierarchy. This is where the "archetype" model of beauty proves misleading. The advantages of beauty do not flow to those who are beautiful in some absolute sense, but to those who are more beautiful than those around them. This is what generates the competitive structure—moving up the beauty hierarchy means bumping someone else down.

None of this would be a problem if people were unable to amplify their natural endowments. Unfortunately, cosmetics and plastic surgery make it possible to reproduce synthetically some of the characteristics that are considered beautiful. As a result, people have the ability to buy their way up the hierarchy. This generates a classic collective action problem. Consider the example of face-lifts. Many women seek to make themselves look younger through artificial means. Unfortunately, how old a person looks is entirely relative. If a woman "looks fifty," it is only because, when compared to other fifty-year-old women, she looks about the same. This means that when a fifty-year-old woman gets a face-lift that makes her look forty, the action can be described in one of two ways. In a sense, she has made herself look younger. But in another sense, all she has done is make all the *other* fifty-year-old women in the population look a little older. These women may then be motivated to get a face-lift *just to retain position*. If this leads all fifty-year-old women to go out and get face-lifts, then their behavior has become perfectly self-defeating. They will be right back where they started—all looking like fifty-year-old women—except that now they will be paying a lot of money to look that way.

This is clearly the dynamic at work in a number of different areas (as any resident of California can attest). Many women would be glad to stop wearing makeup—as long as every other woman stopped too. What they are not willing to do is stop unilaterally, because the *private* cost would outweigh the *private* benefits. It's like standing up to get a better view at a ballgame. You may be able to see better, but only by blocking the person behind you. As a result, once one person stands up, soon everyone else does too. Naturally they would all be more comfortable sitting, and they would be able to see just as well. But sitting down while everyone else stays standing is hardly an option.

Why do women get into this kind of collective action problem and not men? Every so often the cosmetics industry publishes hopeful news bulletins on the “new” body-consciousness among men. Clearly they are hoping that men will begin to engage in similar beauty competition. After all, every arms race is enormously profitable for those who supply the weapons. To date, however, such competition has not materialized. In my view, this is the result of social norms that prevent men from doing so. Use of cosmetics is stigmatized among men as a sign of effeminacy. Much of the power of this norm stems from male homophobia. This is reflected in the fact that beauty competition among gay men is much more intense than it is among heterosexuals, and that critics in the gay community routinely lament the cult of body image that has developed, not to mention the amount of time they have to waste in the gym just to stay in the game.

3 Trust

Collective action problems are an example of how individuals can do things that are against the interests of a group to which they belong because it is in their individual interest to do so. A more subtle class of problems involves the role of trust in stabilizing interpersonal relations. In a collective action problem, people get into trouble because they are acting instrumentally. This is the predominant mode of interaction whenever there is a very large group, so that individual actions are only partially observable, or else in an interaction that is anonymous, so that individuals have no opportunity to communicate or coordinate their actions. On the other hand, in smaller groups or in face-to-face interactions, people’s behavior tends to be more sharply constrained by social norms, and so they are less likely to engage in purely instrumental deliberation. In such contexts, intractable collective action problems are unlikely to arise. The fact that the interaction is normatively constrained, however, creates the opportunity for a new type of “perverse” behavior pattern to emerge—individuals may choose to respect and even enforce norms that are clearly unfair to them. Thus they may actively participate in reproducing institutions that oppress them. The question then is why they would be willing to do so.

This question has become increasingly pressing in recent years, particularly because some of the most spectacularly offensive practices, such as female genital mutilation, are often reproduced by the victims of that same practice. Various accounts have suggested that, even though the practice is expressly designed to secure patriarchal authority, it is often perpetrated by mothers against their daughters. Again, it is tempting to think that there is some kind of ideology at work here—that these women have been brainwashed by men, and so on. This imputation has in turn generated the predictable backlash, in which people point out that the women who do this are perfectly intelligent and reflexive; that they can run off a whole list of reasons why they consider the practice to be justifiable, and so on. Some liberals have even adopted an apologetic stance toward the practice, merely because they don’t like the paternalistic tone that Western critics have taken toward it.¹⁴

Closer attention to the way that stable cooperative practices are maintained helps to shed light on how unjust institutions can be reproduced. By directly prescribing particular patterns of action, social norms permit agents to achieve outcomes that would not be the equilibria of purely strategic interactions. Thus agents stuck in a collective action problem might agree to a rule that constrains their future conduct. The suspects in the prisoner’s dilemma story, for instance, might promise not to turn each other in, or they might become members of a criminal syndicate with a code of silence, and so forth. Norms of this type give each suspect a reason not to perform the action that generates the suboptimal equilibrium.¹⁵ However, even if this reason is one that each of them finds motivating, this is still not enough to secure cooperation. They must also each believe that the other will find the reason motivating, and so will cooperate. One might say that their norm-conformative dispositions must be common knowledge, or more plainly, that they must trust one another.

Unfortunately, a norm-conformative disposition is something that is difficult to just reveal. No matter how much you say you intend to keep your promise, you cannot prove that you intend to, short of actually doing it. This means that people often wind up in situations where they have an opportunity to engage in mutually beneficial cooperation, but where they are not sure they can trust one another.

This is especially problematic where the costs of being exploited are very high, and so the cooperative gains may not be worth the risk. One standard strategy in such a situation is to “build up” trust. People start out by engaging in joint projects in which there is an opportunity for defection, but where the costs associated with being exploited are very low. If neither one takes advantage of this situation, then they may try trusting each other with something more significant. Eventually they reach a point at which future defection would be inconsistent with past behavior (i.e., each can say, “If he was going to sell me out, he would have done it already”).

As a result, small symbolic gestures often play an important role in establishing trust among individuals who do not know each other well. A good example of this is table manners. Although not terribly important in their own right, table manners allow individuals to demonstrate their capacity for self-restraint. Waiting until others are served before starting to eat, for example, shows that one is able to subordinate one’s desires to the requirements of a more abstract social rule. This signals to those around you that you are capable of playing by the rules, and that you intend to do so. This has the effect of putting others at ease, by suggesting that you are unlikely to violate more significant social norms. It tells them that you are not going to do anything unexpected, anything that would make them feel uncomfortable, or take advantage of them.

The collection of these symbolic practices defines what it is to act “normally.” As a result, “being normal,” or “ordinary,” is a normative achievement, one that individuals actively strive for. Someone who behaves abnormally is not only unpredictable, but also untrustworthy. Thus stable cooperation requires that agents, as Terrence Kelly puts it “do being ordinary.”¹⁶ The normality requirement is amplified in situations that require high-trust relations among people who do not know each other well. As a result, these environments tend to have a very standardized appearance and protocol (Kelly refers to this as “hypernormalization”). For example, the usual trappings of a doctor’s office—nutritional posters, white overcoats, eye charts, pastel walls—are there to put the patient in a surrounding that will be immediately familiar, the “doctor’s office.” Everyone knows what a “doctor’s office” looks like. Deviation from the pattern would suggest

that things are somehow not “normal,” and this would immediately cast suspicion on the doctor.

So when agents want to establish their trustworthiness, the best way for them to do so is to “do being normal.” There are a variety of reasons why people want to be trusted. In part, it is because being trustworthy gives them access to cooperative arrangements that can be enormously beneficial to them. But it is also experienced by many as a straightforward moral obligation. Being normal puts other people at ease, it makes them comfortable, allows them to enjoy themselves, let down their guard, and so on. By following the rules, one makes others happy. Violating the rules, on the other hand, immediately raises questions about one’s motives. People have no way of telling whether someone is breaking the rules because she has some kind of principled objection to them, or is rationalizing her self-interest, or is acting out some antisocial tendencies. As a result, it is one thing to criticize a particular social norm in the abstract, it is something else entirely to start violating it in practice. People may well agree that a given practice is unfair to them, that it is exploitative, and so on. But acting on this opinion may involve sacrificing social relationships that are highly valued, for either personal or instrumental reasons. As a result, even people who reject a particular social norm, and who recognize that it treats them, or some other social group, unfairly, may still choose to “play along” with it.

From an instrumental perspective, the advantage of being normal is that it provides access to cooperation. Others will treat you as a “team player,” and not as someone who likes to “rock the boat.” In extreme cases those who don’t play along will be ostracized from the group or community (this is the case with women who are not “circumcised” in societies in which this is the norm—they are often unmarried). There are also somewhat more indirect advantages to securing people’s trust. For example, participating in “shop talk” is extremely valuable to anyone seeking to advance his or her career. In a male-dominated corporation, women who challenge prevailing practices and the ambient corporate culture too directly will often exclude themselves from shop talk with male colleagues. This may not take the form of direct ostracism. Shop talk occurs when people are relaxed, when they’re whiling away time in the corridors or

unwinding in the hotel bar. If a man suspects that something he says may be “used against him” by a female coworker, for example, as part of a sexual harassment grievance, then he will not be relaxed in her company. Women who understand the disadvantage this places them in will often make deliberate use of profanity or sexist language, precisely to signal to the men around them that they will not fault them for doing the same. A woman can establish herself as “one of the boys” by conforming to the rules of a corporate culture that demeans her own gender. Doing so may help her to gain the trust of her coworkers.

The other major advantage of “being normal” is that it reduces the level of anxiety of those around you. Despite the fact that people often complain about rules, most people find normatively unregulated interactions highly vexatious. Unfortunately, when a particular set of rules is challenged, it is seldom the case that a new set of rules springs right up to take their place. It usually takes a lot of negotiation and a certain length of time before people begin to settle into a new pattern—before a new pattern becomes “normalized.” In cases where there is less than universal agreement about what this new pattern should be—and this is most cases—it can be a long time before a new set of rules becomes entrenched. In the meantime, people are left to negotiate their interactions in an ad hoc manner, without the benefit of settled expectations. Most people find this to be, if not intensely anxiety provoking, at least highly demanding. It creates enormous room for misunderstanding, simply because no one knows exactly what the other believes it is appropriate to do, and so cannot tell if the other is trying to “send him a message” by either living up to, or failing to live up to some expectation. Normatively unregulated interaction also requires more interpretive effort, and so more time and attention, both of which are intrinsically scarce resources.

As a result, people often resent those who challenge the rules, even when they do not derive any particular benefit from the terms of the existing institutional arrangements. The benefit they receive stems from the mere *existence* of an established institutional arrangement. Having settled expectations allows them to “get on with it,” without having to worry about the details. This creates an inherent resistance to the problematization of interaction patterns. It can also lead to the

characteristic “backlash” phenomenon that accompanies attempts to effect social change. If an old set of rules becomes discredited, but a new set of rules does not spring up quickly enough to replace them, people may start to revert to the old ones, simply because bad rules are better than no rules.

Consider the example of the controversy over the book, appropriately named *The Rules*.¹⁷ This was a handbook of rules, aimed at a female audience, specifying how to go about courtship, dating, and marriage. Many social commentators were distressed by the retrograde character of these rules, along with their enormous popularity. The rules basically encouraged women to adopt the traditional passive feminine role, to “let him chase you,” make him pay for your meals, and so forth. The reason these rules became popular, however, was not that everyone found them appealing, but simply that some rules were better than none. While the sexual revolution of the 1960s and 1970s succeeded quite well in destroying all of the traditional relationship protocols, it failed to produce many new ones. The absence of such rules turned out to produce a lot more anxiety than freedom. Thus the sexual revolution imposed significant inconveniences upon the generation that came of age in the late 1970s and 1980s, who were largely left without cultural norms to guide them through the difficult transition to sexual maturity.¹⁸

These examples show how people can choose to follow rules that they do not endorse, and so reproduce practices that it is ultimately not in their “interest” to retain. The fact is that cooperation offers significant advantages, and so securing cooperation under unfavorable terms is often superior to securing no cooperation at all. Furthermore, attempts to change social institutions can generate anomic conditions, which lead to failures of cooperation or coordination. Individuals may retain social practices, or seek to restore discredited ones, simply because replacement institutions have failed to materialize. Again, having a bad set of rules is often better than having no rules. Finally, it is an important component of all socialization practices that individuals come to feel some responsibility for the comfort level of those with whom they interact. Opposing entrenched social practices may generate anxiety, tension, misunderstanding, and even hostility in one’s interaction partners. Often the discomfort that

norm-violations cause in others, including significant others, may lead agents to decide that their broader social objectives are simply “not worth it.”

4 Adapted Preferences

The final, and perhaps trickiest, set of problems involves the phenomenon that social theorists refer to as *adapted preferences*. Most people’s goals and desires are strongly influenced by their social environment. The most general reason is that desires are propositional attitudes, and so the range of desires that one can have is internally connected to the vocabulary through which these desires are expressed.¹⁹ Similarly, the range of outcomes that we can imagine achieving are constructed using a set of ideals, roles, and scenarios provided by our culture. Our culture provides the horizon within which we plan our life-projects. As a result, the kinds of things that we want are strongly influenced by what those around us happen to want, or encourage us to want.

It is sometimes felt that this “social construction” of preference creates an opportunity for individuals to exercise undue influence over one another. For instance, critics often complain that the advertising industry cultivates inauthentic desires in consumers, allowing companies to sell people goods that they don’t “really” need. But this kind of criticism must be advanced with care. Apart from its inherently paternalistic structure—the implicit claim that people are not the best judge of their own needs—this argument trades on a distinction between authentic and inauthentic desires that is far from clear. Every desire beyond the most primitive physical urge is socially constructed, and so no *particular* class of them can be discounted by virtue of that fact. The desire to drink single-malt Scotch and read nineteenth-century Russian literature is just as artificial as the desire to drink Budweiser and wear Nike shoes. Most of what passes for “culture criticism” under this guise is just the expression of class distinction.²⁰ If preferences are to be of interest to critical social theory, they must be handled at a more theoretically sophisticated level.

What the concern over adapted preferences points to is not the mere fact that preferences are culturally endogenous, but that they

are also *ambition-sensitive*. In order to protect themselves against disappointment, people usually try to develop “moderate” or “realistic” expectations. In particular, parents often encourage their children to cultivate desires that they have some chance of satisfying—to set their sights high, but not too high. As a result, the goals and desires that people wind up with are usually constrained by what they or others consider it possible for someone in their situation to achieve, or what it is reasonable to expect. As a result, many people who come from disadvantaged backgrounds have downwardly adapted expectations.²¹ They do not want a lot, because they could not reasonably have expected a lot.

The problem with preferences of this type is that they can serve to reproduce inequality, and the effects of discrimination, even after the institutional barriers to advancement have been lifted. So while the first generation of a particular social group might be forced to occupy a particular role, or engage in particular activities, subsequent generations may continue to occupy that role simply because they have learned to like it. In particular, since parents often serve as role models, their children may base their aspirations and expectations on the kind of life that their parents have led. This would be the case, for instance, with women who seek to reproduce patriarchal family structures out of respect for their own mothers, or simply because their own sense of what is valuable in life is tied up with fulfillment of traditional gender roles.

The second major consequence of downwardly adapted preferences is that they may lead members of disadvantaged groups to be too easily satisfied. In a bargaining situation, people generally start out by presenting high initial demands, with the expectation that these will be scaled back as part of the bargaining process. It is in one’s interest to make an initial demand that is quite high, because this leaves more room for concessions. However, it is very important that the initial demand not be *unreasonably* high, since this usually signals a lack of seriousness or good faith. As a result, people with poor background expectations may start out with demands that are in fact much lower than others would be willing to entertain. (For example, people with little work experience generally make salary demands that are too low.) They will wind up with outcomes that are

much worse than those that could be obtained by someone with higher initial expectations. They may also give up on their claims far too easily. One of the most commonly noted characteristics of the upper classes is their extraordinary sense of entitlement. For example, the rich often get a higher level of service simply because they alone have the presumption to demand it. As a result, social inequalities can be projected into the future simply because people have adapted their expectations in response to such inequalities in the past.

The other major problem with preference adaptation occurs when individuals are trying to break out of a particular social role, but are unwilling or unable to break with some of the preference patterns that this role has induced. This can be called the problem of *partially adapted preferences*. In the typical case, an individual has decided that he or she does not want to occupy a particular social role, but retains some collateral preferences associated with it. In the milder cases, this can simply impose welfare losses on that individual. In the extreme, it can actually have the effect of dragging the individual back into a social role that he or she does not want to occupy.

Examples of this phenomenon can be seen in the adjustments that women must make in order to enter the workforce. A lot of attention has been paid to the "second shift" that working women perform.²² The explanation for this is usually thought to reside in the fact that traditional gender norms define certain jobs as "women's work" and thus assign much of the domestic labor to women. This assumes, however, that couples blindly "act out" the roles that their culture assigns to them. In most relationships, however, the division of household labor is a source of considerable conflict. Rather than "play along" with some standard set of rules, couples usually develop a long-run equilibrium through a combination of strategic action and bargaining. Some recent feminist theorists have argued that part of the explanation for the second shift is that women's own preferences sometimes disadvantage them in this bargaining process.²³

Take, for example, the issue of household cleanliness. Whatever the reason, men often have a higher tolerance for filth than women. It is a general characteristic of bargaining scenarios that whoever can "hold out" the longest has a strategic advantage over the other.

Whenever there is a mess to be cleaned up, and no settled agreement on who is to do it, a conflict of interest arises. The person who cleans up does all the work, but then both parties enjoy the satisfaction of living in a clean home. Because of this free-rider incentive, no one will move immediately to clean up, because both will be hoping that the other will do it. But as the house gets messier and messier, the cost of this bargaining stance for each individual increases as well. Eventually, someone will "break down" and do it. The person who experiences the greatest welfare loss from the messiness will generally be the first to break down. Since this is usually the woman, men in heterosexual relationships are often in a strong bargaining position when it comes to cleaning, since both sides know that in a war of attrition, the man can probably hold out longer, and at a lower psychological cost.

Part of the reason for this difference between men and women has to do with women's own values. In the age when women were expected to engage in full-time housework, a woman's social status was intimately connected with how she managed the home—and in particular, how clean she kept it. A filthy home reflected poorly on the wife, not the husband. This has resulted in many women feeling responsible for housework, even as they explicitly renounce this social role. For example, many women who work full-time employ cleaning services. Despite this fact, it is not unheard of for women to "pre-clean" the house before the maid arrives (a compulsion that men generally find baffling). This is an example of how some individuals, despite having officially renounced a particular social role, may not be fully adapted to its renunciation.

Inadequately adapted preferences can also generate unanticipated consequences. For example, in the traditional patriarchal "marriage market," a woman's primary asset was her beauty, whereas for a man it was his wealth. Many people look to marriage as a way of enhancing their social status. Among men, this has meant intense competition for beautiful women, and among women, competition for men with significant wealth or earning potential. There is evidence that this preference pattern has persisted, despite the mass entry of women into the labor force. For example, a study of Stanford MBA graduates showed that men earned an average salary of \$144,461,

while women earned \$101,204. Here one can see a typical "gender gap" in earnings. Even more startling, however, is the gap in what their spouses earned. The female MBA's husbands earned \$120,124 on average, while the average male MBA's wife earned only \$30,323.²⁴ The message is fairly clear—despite making extremely high salaries, most of the female MBAs in this study still managed to marry men who earned more than they did.

The full explanation for this type of marriage pattern must obviously involve a complex set of factors (which will include the gender gap, along with the fact that women tend to marry men who are older than they are). According to Rhona Mahoney, however, one important factor that must be taken into consideration is straightforward preference:

I suspect that many hard-driving women believe so strongly in the value of hard work in a challenging field that a man who takes a more relaxed approach may seem second-rate. A man who earns very little, because he is a community organizer or struggling writer, may even seem lazy or suspect. For example, in one study of how sixty-three women made decisions about their careers and their marriages, most women said they were not interested in marrying a househusband. The women who were very devoted to their paying work wanted a man who was, too.²⁵

This preference structure, where it occurs, can have unanticipated consequences. It is hard for women with careers—who must take specific evasive action to avoid becoming housewives—to resist the tendency to *devalue* housework. As a result, they do not want to marry househusbands, precisely because they want to marry someone whose life-choices they respect. But marrying someone who makes more money than you is tantamount to marrying someone with more power than you. This kind of power asymmetry may not become salient until many years into the relationship. For example, when child-rearing responsibilities begin to exert time pressure on the household, the person most likely to cut back on work responsibilities is the person with the lower salary. As a result, women who show preference for men with "high-powered" careers have a preference structure that is often not fully adapted to their own ambitions.

The underlying phenomenon here simply reflects the fact that social change can often be very complicated. Social practices are sup-

ported by a vast network of behavior and preference, whose functional connection to the reproduction of any particular practice may be very difficult to discern *ex ante*. This means that changing a practice may involve far more extensive revisions in individuals' preferences than anyone ever anticipated. As a result, a social practice may persist despite enjoying very little expressed support. This may be because individuals remain attached to preferences and behavior patterns that indirectly support it, or that reproduce it through consequences that begin to be felt only in the very long term.

The general goal of this paper has been to examine why individuals sometimes act in ways that are contrary to their own interests. The problem of adapted preferences arises when people have *interests* that are, in some sense, contrary to their interests. But in developing a critical analysis of this problem, I have resisted the temptation to appeal to some external conception of what individuals' "real" interests are. This kind of high-handed—not to mention epistemologically suspect—maneuver is by now thoroughly discredited. I have focused on cases where the preferences that individuals have are in tension with one another (partial adaptation), or else cases where the way that these preferences interact with those of others generates lower overall levels of preference-satisfaction (downward adaptation). Neither of these critical strategies requires any second-guessing of the agent's own self-conception, or involves any suggestion that the agent is somehow irrational for failing to recognize where her "real" interests lie.

5 Conclusion

The explicitly stated goal of this paper has been to wean social critics from their attachment to the concept of ideology. The general concern is that, through an excessively uncharitable attitude toward their subjects, critical theorists have had a tendency to undermine the credibility of their own views. In the background, however, has been another concern. Many social critics succumb to a sort of tacit cultural determinism. This is reflected in the widespread assumption that social practices directly reflect people's values, or that they express some set of beliefs about how people should act. If this were the

case, then the key to changing social institutions would indeed be to change people's values or beliefs. Unfortunately, while some social practices are directly "patterned" by the cultural system, many more are reproduced through very loosely constrained strategic action. These interactions are integrated only indirectly, and so the associated outcomes may not reflect any specific set of values or beliefs. In this case, social criticism alone will not change anything.

The more serious problem for critical theory arises as follows: after having presented the criticism, and having it widely accepted, the critic expects to see some kind of social change. When none is forthcoming, the critic begins to suspect, not that there is a practical problem preventing implementation of the desired improvements, but that the criticism itself was too superficial, that it didn't get to the root of the problem. The ideology must be more pervasive than originally suspected. Perhaps the original criticism was insufficiently radical, because it used concepts that were in general circulation, and hence complicit in the ideological system. The solution may be to deconstruct these concepts, and form an entirely new set.

Once this line of thinking has been engaged, the critical theory becomes increasingly baroque, increasingly obscure, and of course, increasingly unlikely to change anything. This can generate a vicious cycle of theoretical self-radicalization, in which critics respond to the increasing irrelevance of their theories by further radicalizing them, making the entire apparatus more and more remote from the concerns and the vocabulary of everyday life. The goal of this paper has been to suggest one way in which critical theorists can engage in social criticism without pricing themselves out of the market. More attention to the structure of social interaction, the practical mechanisms through which undesirable interaction patterns are reproduced, has the potential to generate more useful theoretical interventions.

Notes

1. Ludwig Feuerbach: "Every limitation of the reason, or in general the nature of man rests on a delusion, an error." *The Essence of Christianity*, trans. George Eliot (New York: Harper, 1957), 7. Karl Marx: "In all ideology men and their circumstances appear upside-down as in a camera obscura." *The German Ideology* (Moscow: Progress Publishers, 1976), 42.

2. Donald Davidson, "On the Very Idea of a Conceptual Scheme," *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1984), 183–198.
3. Lucien Lévy-Bruhl, *How Natives Think*, trans. L. A. Clare (London: Allen, 1926).
4. E.g. Sigmund Freud: "Religion would thus be the universal obsessional neurosis of humanity," *The Future of an Illusion*, trans. W.D. Robson-Scott (New York: Anchor 1961), 70–71.
5. Karl Mannheim, *Ideology and Utopia*, trans L. Wirth and E. Shils (New York: Harvest, 1936), 75–83.
6. Freud's inner circle often conducted theoretical disputes by analyzing each other's motives for disagreement. See Peter Gay, *Freud: A Life for Our Time* (New York: W.W. Norton, 1988).
7. Figures reported by Edward E. Zajac, *The Political Economy of Fairness* (Cambridge: MIT Press, 1995), 29.
8. See Allen Buchanan, "Revolutionary Motivation and Rationality," in *Marx, Justice and History*, ed. M. Cohen, T. Nagel, and T. Scanlon (Princeton: Princeton University Press, 1979), 264–287.
9. Mancur Olson, *The Logic of Collective Action* (Cambridge: Harvard University Press, 1965).
10. These figures and calculations are from Naomi Wolf, *The Beauty Myth* (London: Vintage, 1991), 88–89.
11. *Ibid.*, p. 83.
12. *Ibid.*, 6.
13. *Ibid.*, 87. Here Wolf is conforming to one of the stranger unwritten rules of twentieth-century social criticism, namely, that it is okay to call people stupid as long as you blame someone else for their stupidity.
14. See Yael Tamir, "Hands off Clitoridectomy," *Boston Review* 21/3 (1996).
15. For defense of this claim see Joseph Heath, "Foundationalism and Practical Reason," *Mind* 106 (1997): 451–473.
16. Terrence M. Kelly, "Rationality, Reflexivity and Agency in the Critique of Everyday Life" (unpublished Ph.D. diss., St. Louis University, 1998). Much of this discussion is indebted to Kelly's analysis.
17. Ellen Fein and Sherrie Schneider, *The Rules: Time Tested Secrets to Capture the Heart of Mr. Right* (New York: Warner, 1996), and their follow-up, *The Rules II: Living the Rules* (New York: Warner, 1998).
18. A situation memorably dramatized in the film "The Ice Storm." This is in part what motivates the often noted "ironic" disposition that characterizes members of this generation. Irony allows people to follow "old" rules without being taken to endorse them.

19. Charles Taylor, "Interpretation and the Sciences of Man," *Philosophy and the Human Sciences: Philosophical Papers 2* (Cambridge: Cambridge University Press, 1985), 15-57.
20. See Pierre Bourdieu, *Distinction*, trans. R. Nice (Cambridge: Harvard University Press, 1984).
21. See Jon Elster, *Ulysses and the Sirens* (Cambridge: Cambridge University Press, 1979). See also G. A. Cohen, "Equality of What? On Welfare, Goods, and Capabilities," in *The Quality of Life*, ed. M. Nussbaum and A. Sen (Oxford: Clarendon Press, 1993), 9-29.
22. Arlie Hochschild, *The Second Shift: Working Parents and the Revolution at Home* (New York: Viking, 1989).
23. See Rhona Mahoney, *Kidding Ourselves* (New York: Basic Books, 1995).
24. For further discussion, see Mahoney, 143-145.
25. *Ibid.*, 146. The study referred to is from Kathleen Gerson, *Hard Choices: How Women Decide about Work, Career, and Motherhood* (Berkeley: University of California Press, 1985).