

# “False Consciousness” and Democracy

Matthew McCoy (msmccoy@princeton.edu)  
Princeton University

Prepared for the Annual Meeting of the WPSA  
April 2014

This is a Draft. Please do not cite or circulate.

## **1. Introduction**

This paper presents a critique and qualified defense of the much maligned concept of false consciousness. A critique of a concept is, as Elizabeth Anderson puts it, “not a rejection of that concept, but an exploration of its various meanings and limitations.”<sup>1</sup> The goal of critique is to replace one ill-defined and overburdened concept with distinct “ideas that we can choose to use or not, depending on how well the use of these ideas suits our investigative purpose.”<sup>2</sup> As I hope to show, the concept of false consciousness is long overdue for this kind of critical analysis. The term “false consciousness” is used to refer to superficially similar phenomena that are in fact marked by some significant differences. Uncovering these differences, I tease apart two ideas of false consciousness that tend to be conflated in debates about the concept and argue that one of these two ideas deserves a place in the conceptual toolkit that we, as members of democratic society, use to evaluate and criticize the behavior of our fellow citizens.

## **2. The Controversy over “False Consciousness”**

Let me start with a provisional definition of false consciousness and say something about why the concept, understood in these terms, has been the subject of persistent controversy. To charge someone with false consciousness is to charge him with failing to recognize and act on his own interests.<sup>3</sup> False consciousness thus implies both cognitive error (a failure to recognize some facts about one’s interests) and motivational error (the desire to act in a way that frustrates one’s interests). This is why we speak of false consciousness—where “consciousness” refers to a complex of beliefs, desires, and attitudes—rather than simply “false beliefs” or “false desires.”<sup>4</sup> False consciousness is what we might call a diagnostic concept, in that it serves to identify a certain defect in an individual’s politically relevant beliefs and desires—namely, their incompatibility with his or her interests. Over the years, the concept has been developed and

deployed by those concerned to expose unseen forces that act to sustain putatively illegitimate asymmetries of power. Theorists of false consciousness tend to share the concern that when members of subordinated groups fail to recognize their own interests, they are likely to become unknowingly complicit in their own mistreatment, reproducing practices and institutions under which they suffer.<sup>5</sup> Identifying instances of false consciousness is thus seen as a means of exposing a particularly insidious way in which asymmetries of power are perpetuated over time.

Though the concept of false consciousness has its roots in radical social theory, it has made its way into our mainstream political discourse, where it's been taken up most notably by analysts of democratic politics. Today we frequently encounter variations of the observation that "people are voting against their interests" from commentators both inside and outside the academy.<sup>6</sup> Sometimes these observations are made with explicit references to "false consciousness," but even when they are not, they implicitly affirm the concept's underlying tenets—namely, that there is a concrete standard of interests against which individuals' political preferences can be judged and that this standard is knowable to outside observers. It's not hard to see why those interested in democratic dysfunction would take an interest in the problem of false consciousness. Democracy empowers the people themselves, through more or less direct means, to decide how they will be governed. If what the people choose for themselves reflects a misunderstanding of their own interests, we might think, there is reason to question the outcomes of certain forms of democratic decision-making.<sup>7</sup>

But for every commentator who employs the language of false consciousness, there are those who insist that the practice of attributing false consciousness to others is "highly objectionable."<sup>8</sup> Though critics object to talk of false consciousness for a variety of reasons,<sup>9</sup> the most common and, in some sense, most fundamental reason for objecting to talk of false

consciousness stems from the thought that someone who engages in attributions of false consciousness lacks justification for claiming to know the true interests of the targets of his criticism.<sup>10</sup> The essence of the objection is nicely captured by Rosen's question: "What entitles [observers of false consciousness] to assume that they stand on the vantage-point of truth and enlightenment so that they know better than the individual agent in question what consciousness the agent ought to have?"<sup>11</sup> To the extent that the observer of false consciousness can't supply a persuasive answer to this question, his charges register as little more than attempts to discredit his political opponents. Rather than accept that others might have good reasons for seeing things differently than he does, it appears, the observer of false consciousness simply concludes that those with whom he disagrees must be in the grips of some sort of delusion. This way of dealing with disagreement would be objectionable in many contexts, but it appears especially odious in the context of democratic politics, where we expect disagreement to be taken seriously, if not treated respectfully, and not dismissed as a product of the other side's irrationality.

In what follows, I aim to clarify the justificatory burden raised by charges of false consciousness and ask whether that burden can be met in a way that doesn't offend against democratic values. I will, however, approach this goal obliquely, beginning from the observation that there are a couple of distinct ideas that get lumped together under the heading of "false consciousness." Only after teasing these ideas apart do I attempt to determine whether either can answer to critics' demands for justification. The paper thus undertakes two relatively discrete tasks. The critique of false consciousness that I refer to in the introduction comprises sections 3 and 4 of the paper. Section 3 distinguishes several senses of "interest," a term which, while clearly central to debates about false consciousness, often remains poorly defined by parties to those debates. Relying on this typology of interests, section 4 draws a distinction between two

senses in which someone might be said to mistake his own interests and uses this distinction to argue that we ought to think of the unitary concept of false consciousness as dividing into two separate ideas. In the second half of the paper, sections 5 and 6 evaluate these two ideas of false consciousness and ask whether either can meet the justificatory burden it raises. Ultimately, I argue that while both ideas of false consciousness can be justified in certain contexts, only one of the two ideas furnishes a worthwhile tool for social and political criticism within a democratic polity.

### **3. “Interest” and its Meanings**

When defenders and critics of false consciousness clash over the questions of whether and when we are justified in charging someone with mistaking his own interests, they are often operating under different ideas of just what an “interest” is. The confusion is not entirely surprising. Scholars have long noted the “diverse and shifting manner” in which the term “interest” is used in our political discourse.<sup>12</sup> Jane Mansbridge goes as far as suggesting that it may be impossible to define the term adequately.<sup>13</sup> That said, we can certainly improve upon the relatively undisciplined and imprecise usage of the term we tend to see in debates about false consciousness. One way to begin carving up the conceptual space occupied by our notion of “interest” is to treat that notion as “bifurcated into two separate senses,” each with multiple, more or less distinct variants.<sup>14</sup> On one hand, “interest” can refer to the objective relation of being concerned or invested in something with respect to one’s advantage or disadvantage. On the other hand, “interest” can refer to a subjective feeling of concern or curiosity.<sup>15</sup> In this section, I want to explore these two senses of interest and some of their variants, creating a typology of interests that will orient the argument in the subsequent sections of the paper and

prevent against the sort of terminological ambiguities that derail so many discussions of false consciousness.

Let's begin with the objective sense of interest. [On one variant of the objective sense, saying that "A has an interest in X" is roughly equivalent to saying that "A has a stake in X." This captures both technical-legal uses of "interest" like "Mary has an interest in a software company," as well as more general uses like "the residents of Cleveland have an interest in the upcoming mayoral election." Because the existence of these sorts of interest relations is not directly dependent on the mental states of the affected individuals, one needn't have any particular knowledge of or attitude toward her interests in order for them to qualify as her interests. Though people often know and care about the things in which they have a stake, it is possible that a person might be wholly unaware that she has an interest in something, just as it is possible that a person might recognize her interest in something while remaining ambivalent about that thing. Take the example of Cleveland's mayoral election. Some residents of the city might not know about the election, while others might know about the election but remain entirely unconcerned about its outcome. Nonetheless, to the extent that these men and women will be materially affected by the policies of the future mayor, they can be said to have an interest in the election, a stake in its outcome. Because objective interests are not directly dependent on individuals' beliefs and attitudes, they can be "observed and measured by standards external to the individual's [or individuals'] consciousness."<sup>16</sup> This means that an outside observer may find herself in a position to say what the objective interests of others are, and perhaps even to challenge those others' own beliefs about their interests. This is not to say that ascriptions of objective interests are always straightforward and uncontroversial. But where there is a question of whether or not "A has an interest in X," in the manner I've been describing,

it is a question of whether or not X affects A in a certain way, not a question about what A knows or thinks about X.]

On the second variant of the objective sense, “A has an interest in X” means not simply that “A has a stake in X,” but that “A is benefited by X” or that “X is good for A.” This is how we use the term “interest” when we say, for instance, that children have an interest in a calcium-rich diet or that men and women of a certain age have an interest in regular exercise.<sup>17</sup> This same idea can, of course, be expressed with slightly different phrasing. Rather than “children have an interest in a calcium-rich diet” we sometimes say “it is in the interest of children to eat a calcium-rich diet.” The interests captured by this use of the term are, like those in the previous examples, objective in that they do not directly depend on the mental states of the agents to whom they are ascribed: a child who wants to eat only candy still has an interest in a calcium-rich diet. It is also possible, at least in principle, for an outside observer to know what is in the interest of (i.e., good for) another person without inquiring into that person’s thoughts or feelings. Again, one needn’t know what a child likes to eat in order to know what sort of diet is in his interest. However, it bears noting that when we use “interest” to describe what benefits a person, we take on a more substantial burden of proof than we do when we use “interest” simply to describe what sorts of things a person has a stake in. Consider the difference between saying that the residents of Cleveland have an interest in the upcoming mayoral election and saying that it is in the interest of Cleveland residents to vote for the Democrat in the upcoming mayoral election. The latter claim involves a number of implicit judgments that the former does not, each of which might become the subject of disagreement. Let me briefly examine a few such judgments.

[If there are certain things that are good for all persons *qua* persons, it follows that there will be interests that are universally shared, since all persons will have an interest in pursuing or having access to those things that are good for them *qua* persons. But it is important to recognize that objective interests are not necessarily universal, since what is good for one person may not be good for another. It may, for instance, be in the interest of a professional cyclist to consume 4,000 calories a day given the physical demands of her sport. This diet's benefit to the cyclist is objective since it does not depend on her adopting a certain attitude toward or belief about it, but the diet's benefit is hardly universal since it would do nothing to help (and, indeed, would likely harm) those engaged in less physically demanding pursuits. Typically, then, the claim that "X has an interest in A" involves a judgment about what type of person A is—that is, a judgment to the effect that X is good for A by virtue of the fact that A is a child, an athlete, a worker, or what have you. Such judgments are often straightforward and uncontroversial, but they can become the subject of disagreement when the categories we use to distinguish a type of person are vague and/or there is some question as to whether a particular person fits into a given category. We may, for instance, disagree about what makes a person a worker for the purpose of evaluating an economic policy advertised as being in the interest of workers.

Additionally, when we claim that "A has an interest in X" we are usually making a judgment that A has some need that X satisfies. Consider again the example of children having an interest in a calcium-rich diet. It is in the interest of children to consume a calcium-rich diet *because* children need calcium to ensure healthy bone development. As William Connolly points out, these sorts of need claims are always at least implicitly triadic: "Person A needs X in order to (do, be, or become) Y."<sup>18</sup> And, as Connolly puts it, "the Y in this triad is always the crucial variable."<sup>19</sup> In some cases, like when we say that children have an interest in a calcium rich diet,



the “Y” in the triad is uncontroversial enough to be taken for granted. This will usually be the case when the “Y” is some basic biological need. Other times, however, the “Y” may require a more explicit defense. For instance, the claim that persons need a certain type of liberal education to enable them autonomously to select and revise their conceptions of the good might be challenged on the grounds that autonomously selecting one’s conception of the good reflects a particular cultural ideal rather than real need. The point is not that a defense of autonomy (or any other need) is impossible to provide, but simply that claiming, “A has an interest in X” depends on a judgment about “what X is needed for.”<sup>20</sup> ]

We make a third type judgment when we talk about what is in a person’s best interest. Strictly speaking, “A has an interest in X” implies only that X is one of the things in which A has an interest. To say that “X is in A’s best interest” or something along these lines is to make one of several stronger claims: that X will harmoniously advance all of A’s interests;<sup>21</sup> that X will advance one or more of A’s interests without affecting any others; or that X is so important to A’s well-being that the setbacks it may cause to his other interests are outweighed by its benefit. Judgments about what is in a person’s best interest are made difficult by the fact that most persons have multiple and competing interests. Feinberg captures this difficulty with the following example: “going alone to Europe to do research for a few months may present unique opportunities for professional achievement to a professor while also creating family strains.”<sup>22</sup> The professor in Feinberg’s example finds himself in the difficult though not unfamiliar position of facing a decision that pits one of his interests against another. Is the research sabbatical in his best interest? Such a determination can proceed only from a judgment as to the weights of the interests at stake.<sup>23</sup> Crucially, negative statements like “A is acting against his interests” also typically involve this sort of judgment. In other words, “A is acting against his interests” does

not normally imply just that A faced a decision that pitted some number of his interests against one another and he inevitably chose a path that advanced one interest at the expense of the others; it implies that A faced a decision in which one option was clearly in his best interests, and he chose incorrectly. This point will be particularly important when it comes to evaluating a certain idea of false consciousness which implies that the target of the charge is acting against her (best) objective interests.

One final issue to consider is that what is “good for” a particular person will be determined, at least in part, by her own purposes in life. Take the example of the professional cyclist. We can say that *qua* cyclist, it is in her objective interest to eat 4,000 calories a day. But one is not born into being a professional cyclist. While it is now in the objective interest of the cyclist to eat a particular diet, she is a cyclist in virtue of her own commitments. In other words, what is in her objective interest now is, at least in part, the product of her own projects and plans—a result of the things in which she has chosen to invest. To say this is not to assume that anything one chooses to invest in becomes valuable (and generates subsidiary interests) by virtue of that investment. It is only to assume that there is a more or less extensive range of valuable projects in which one might choose to invest, and that investing in one rather than another affects what objective interests a person has. This is why it is misleading to say that when we are dealing with objective interests, “[w]hat is in your interests depends not on your desires, nor on any class of them, but on what is the good life for you.”<sup>24</sup> It may be true that wanting something is not sufficient to make it an objective interest. But if “the good life for you” is determined in part by what you desire to do and become, then what is in your interests does depend, albeit indirectly, on a class of desires associated with certain long term goals.<sup>25</sup>

Let's now turn to the subjective sense of "interest." As noted above, the subjective sense of "interest" refers to a feeling of concern or curiosity rather than an objective relation. When we use "interest" in the subjective sense, saying that "A has an interest in X" is more or less equivalent to saying that "A is interested in X." However, we need to distinguish between two different sorts of claims that may be implied by "A is interested in X." When "X" is a thing like Italian films, "A is interested in X" means something like "A enjoys learning about X" or "A finds X interesting." To say that John is interested in Italian films is thus to imply that he enjoys watching them, that he finds them interesting, that he knows something about them, and so on. However, when "X" is an activity like building a house or going to college, "A is interested in X" implies that A is subjectively motivated to do X. Thus, if we say that John is interested in becoming a film director, we mean that he wants to direct films some day. It is this second variant of the subjective sense of "interest" that I want to focus on here.

We saw above that objective interests exist irrespective of persons having any particular knowledge of our attitude toward them. This is not the case with subjective interests. One has subjective interests in activities or projects only insofar as he adopts a certain attitude towards them. It would make little sense to say that someone had a subjective interest that he didn't know or care about. However, we should not conclude from this that persons can never be mistaken with respect to their subjective interests or have desires that conflict with those interests. It is true that we sometimes use the word "interest" in a way that equates interests with mere passing wants. If I ask whether you have any interest in seeing a movie after work, and you say do, it is hard to see how you could be mistaken about that interest. But we have many interests which do not have the immediacy of passing wants. Some of our subjective interests have the character of "relatively deep-rooted and stable" aims.<sup>26</sup> Though we cannot normally be mistaken about what

our subjective interests are, since they would not be our interests if we did not recognize them as such, we can be mistaken about how to advance them. This is especially true in the case of relatively abstract or ulterior interests. I may have a subjective interest in pursuing a career in politics and yet be entirely mistaken in my judgments about how best to achieve that goal. Likewise, even though one's subjective interests are a product of one's wants, it is possible for persons to have wants that conflict, and do so objectively, with their subjective interests. Though I am interested in becoming a successful politician I may desire to do things which set back this interest, either because I wrongly think that the things I do are instrumental to my goal, or because I experience weakness of will and desire to do them in spite of their consequences.<sup>27</sup>

Of course, subjective and objective interests overlap in practice. For one thing, we typically expect people to take a subjective interest in pursuing or having access to those things in which they have an objective interest. In other words, we expect people to be subjectively motivated to pursue things that are objectively good for them. It may even make sense, under the right conditions, to treat subjective interests as a proxy for objective interests. It is also true that subjective interests with the character of "relatively deep-rooted and stable" aims can give rise to subsidiary objective interests. Consider again the example of someone who wants to become a director. Assuming this ambition represents a valuable aim, we might say that given his ambition, it is in his (objective) interest to attend film school. That said, it is important to bear in mind that there is no necessary connection between one's objective and subjective interests. To say that someone has a subjective interest in doing or becoming X does not necessarily imply that X is good for him but only that he has a certain attitude towards X.

With these distinctions in hand, let me now turn to the concept of false consciousness.

#### **4. Two Ideas of False Consciousness**

According to the provisional definition with which we began, to charge someone with false consciousness is to charge him with failing to recognize and act on his own interests. Drawing on the previous section's typology of interests, I want to argue that we ought to think of the unitary concept of false consciousness as encompassing two distinct phenomena. One form of false consciousness, call it deep false consciousness, involves error in the identification of one's objective interests and the formation of desires that fail to aim at one's good. The other form of false consciousness, call it shallow false consciousness, involves error in the identification of what serves one's subjective interests and the formation of desires that are incompatible with one's own plans and purposes. Though both forms of false consciousness fit the provisional definition of false consciousness laid out above, and both are sometimes discussed under the heading of false consciousness, I want to argue that they are distinct in crucial ways.

I begin with the idea of deep false consciousness. Assuming that persons have some objective interests—that is, some things that are objectively good for them either *qua* persons or in virtue of some particular characteristics that they possess—it is possible that a person might err in identifying those interests. One might fail to recognize some number of her interests, or might believe she has an interest where she actually has none. An agent who mistakes her interests in these ways makes what we might call a first order error with respect to her interests, which is to say, an error in determining what sorts of things are good for her and, accordingly, what sorts of ends she ought to pursue. Note that someone who misperceives her interests in this manner might be quite capable of taking effective and efficient means to advance her perceived interests, behaving in a manner that is “purposive, goal-oriented, [and] guided by considerations of instrumentality.”<sup>28</sup> Nor have we any reason to think that those who err in the identification of their interests are incapable of providing “a coherent account of why they are acting as they are,

by reference to what they seek to achieve and how they expect what they are doing to advance their goals.”<sup>29</sup> In other words, first order error does not entail instrumental irrationality or the inability to account coherently for one’s actions.<sup>30</sup>

However, when one fails to recognize her objective interests, her self-regarding desires will go awry in a particular way: they will fail to “aim at her well-being,” as Denise Meyerson puts it, or will aim at her well-being only accidentally.<sup>31</sup> Building on Meyerson’s metaphor, we might say that these desires are errant in the sense that they miss their target, which is one’s own good. It is important to emphasize that the kind of errant desires I’m describing here are a class of self-regarding desires. A desire to sacrifice oneself for the good of another does not aim at one’s well-being, but it does not therefore miss its target. Errant desires are desires for oneself that do not conduce to one’s well-being. Of course, it is possible to have errant desires that do not stem from false beliefs about one’s objective interests. Someone may have a complete understanding of what her objective interests are and yet form desires that fail to aim at those interests if she suffers from weakness of will. Though Meyerson appears to think that errant desires should always be considered a form of false consciousness, I am inclined to think that only those errant desires which stem from a misunderstanding of one’s objective interests ought to be considered constitutive of false consciousness. To be sure, one can define “false consciousness” however she chooses to, but Meyerson’s definition stretches the category of false consciousness beyond the point of usefulness by obscuring some intuitively significant distinctions. On her account, for instance, the individual who chooses to eat a second helping despite knowing he should watch his weight suffers from false consciousness. I think it is important to draw a distinction between this type of individual and the individual who is under some kind of misapprehension about what is and isn’t good for him, so I prefer to stick with a

narrower understanding of which sorts of errant desires constitute false consciousness, though I readily admit that this is not a knockdown argument against Meyerson's position.

To charge someone with deep false consciousness is to engage in a form of external criticism that measures an agent's beliefs about her interests against a true account of her interests furnished by some theory of well-being. The idea of deep false consciousness is thus substantive in the sense that it defines false consciousness, at least implicitly, in opposition to a substantive standard of "correct" consciousness supplied by an underlying account of objective or "real" interests. There are different ways to articulate or conceptualize the idea of deep false consciousness. Indeed, there can be as many unique conceptions of deep false consciousness as there are unique accounts of real interests. In the next section, I'll examine some challenges related to the search for a defensible account of real interests. First, however, I want to distinguish deep false consciousness from what I'm calling shallow false consciousness.

Whereas deep false consciousness involves failure to correctly identify one's objective interests, shallow false consciousness involves failure to recognize how various courses of action serve or set back one's subjective interests, that is, one's own aims and purposes. To understand the sort of failure involved in shallow false consciousness, consider someone who takes himself to have an interest in attaining economic self-sufficiency, but because of flawed reasoning, lack of information, or some combination of the two, fails to see how to attain such a state. He makes the wrong investments, takes the wrong jobs, and thus frustrates his own purposes. His error doesn't involve adopting the wrong end, but pursuing it in the wrong way. We can call this second order error to distinguish it from the sort of error involved in deep false consciousness. Second order error is error in instrumental reasoning, the failure to recognize the appropriate means to a given end rather than error in the selection of an appropriate end. While the example

of the man striving in vain to achieve economic sufficiency presents a case of someone failing to advance a subjective interest that arguably looks like an objective interest as well—in other words, it seems plausible that economic sufficiency is good for him in addition to being something he is subjectively motivated to attain—it is important to recognize that the question of whether or not someone misidentifies the means of advancing a subjective interest is entirely separable from the question of whether or not that subjective interest is aligned with an objective interest. This point is significant, as we will see, for it means that one can remain agnostic on the question of whether an agent’s subjective interests are objectively valid while identifying ways in which that agent’s beliefs about how to advance his subjective interests go awry.<sup>32</sup>

Like first order error, second order error leads to the formation of defective desires, but the defect of the desires is different in the two cases. When someone has false beliefs about how to advance his subjective interests, his desires will be inconsistent with his underlying aims. Suppose that someone has a subjective interest in doing X and wrongly believes that doing Y will advance this interest. Though he desires to do Y, the desire is a distorted expression of his interest in X, predicated on his false belief about the relationship between X and Y. The defect of such a desire is not that it fails to aim at an agent’s good, though it may indeed fail to do so, but that it is unsupported by the reasons the agent has for holding it. Raz has called these sorts of desires “false desires.”<sup>33</sup> But since the term “false” is already overburdened and poorly understood in discussions of false consciousness, I prefer the term distorted desires, which emphasizes the sense in which such desires are distorted expressions of their underlying purpose.

To charge an agent with shallow false consciousness is to engage in a form of internal criticism that measures her beliefs and desires against her own underlying aims and purposes. The criticism is internal because it takes the agent’s own subjective interests as the standard



against which beliefs and desires are measured. Whereas conceptions of deep false consciousness define false consciousness in opposition to a substantive standard derived from a theory of well-being, shallow false consciousness is a purely formal concept. The “falseness” of shallow false consciousness consists in the fact that one’s beliefs and desires are inconsistent with her underlying aims and purposes, not in the fact that they fail to conform to a substantive standard. Another way to express the different implications of deep and shallow false consciousness is to say that while conceptions of deep false consciousness always presuppose something like true or correct consciousness, shallow false consciousness is a purely negative concept. One needn’t have a positive, underlying theory of well-being that generates an account of real interests in order to identify the “falseness” of shallow false consciousness. What one needs is some means of revealing the divergence between an individual’s underlying aims and her beliefs and desires related those aims.

As far as I know, I’m the first to suggest explicitly disaggregating the concept of false consciousness in this way, but the distinction I’ve been developing here has been implicit in some previous discussions of the concept.<sup>34</sup> In his recent “In Defense of ‘False Consciousness,’” Steven Lukes objects to the notion that working class Americans who vote on the basis of social rather than economic issues can, on those grounds alone, be described as suffering from false consciousness. “It is arbitrary,” he says, “to just assume that people’s real interests are economic; we should take their priorities seriously, not assume them to be deluded because they see things differently. If one believes that abortion is murder, how could one reasonably allow it to count less than one’s material comforts, or even necessities?”<sup>35</sup> On the other hand, Lukes suggests that it may be appropriate to attribute false consciousness to a man described by a *New York Times* reporter as living off of Social Security and Medicare while advocating policies which would

jeopardize these programs. Why is the first attribution of false consciousness “arbitrary” while the second is not? Lukes does not provide any sort of systematic answer to that question. My sense, however, is that the key difference between the two cases is that the first involves questioning the priorities that individuals set for themselves while the second involves questioning individuals’ beliefs about how to advance the priorities they are thought to have (though, of course, may be found not to). In that sense, the difference between the two cases closely resembles the difference between what I’ve been calling deep and shallow false consciousness. In the next two sections, by evaluating these two ideas of false consciousness, I will try to show more systematically than Lukes does why judgments about the justifiability of the two sorts of charges would diverge.

## **5. Deep False Consciousness**

Let me begin this section by previewing its conclusion. I want to argue that the idea of deep false consciousness offers little value as a tool of social and political criticism. This is not the same as saying that ascriptions of real interests are necessarily arbitrary, as Lukes might be read to suggest and as others have argued explicitly. Part of what I want to do in this section is show that this common, perhaps intuitively appealing argument for rejecting deep false consciousness overstates the case against the idea. Rather than joining the majority opinion against deep false consciousness, then, I want to issue a kind of concurrence: I think the idea of deep false consciousness is flawed but neither as deeply nor for the reasons it is often taken to be. Part of this concurrence will involve showing the insufficiency of an approach to defending deep false consciousness adopted by some of the idea’s proponents.

Before I lay out the existing arguments on both sides of this debate, I need to expand on the previous section’s brief remarks about the justificatory burden raised by the idea of deep

false consciousness. As noted, deep false consciousness is always defined in opposition to a substantive standard of objective or “real” interests generated by an underlying (often implicit) theory of well-being. Whether or not a particular conception of deep false consciousness can meet the justificatory burden it raises will depend on whether its account of real interests is defensible. I’ll refer to this as the defensibility requirement. The term “defensible” is admittedly somewhat vague, because what counts as a defensible account of interests will depend on background norms that may vary from context to context. But since the purpose of the present investigation is to determine whether some idea of false consciousness deserves a place in the conceptual toolkit that we, as members of a democratic society, use to evaluate the political behavior of our fellow citizens, I’ll interpret “defensible” as meaning defensible within a society whose members endorse an idea of democratic toleration that enjoins respect for the reasonable diversity of citizens’ values, projects, and purposes in life.

This idea of toleration may seem to incline in favor of a more minimalist, liberal understanding of interests and, thus, unfairly bias the analysis towards the conclusion that most conceptions of deep false consciousness will fail to meet the justificatory burden they raise. Let me offer a two-part response to this worry. First, I am willing to grant that my argument here will do little to persuade the avowed anti-liberal who sees no need respect the reasonable diversity of values. The question I want to ask in this paper is, given a background of basic democratic values, what sorts of charges of false consciousness, if any, may be justified? However—and this is the second part of my response—it is worth noting that the idea of democratic toleration I lay out above, though it may have a liberal flavor, has come to have a fairly ecumenical appeal. In particular, it has been largely taken up by political theorists on the left, including those working to rehabilitate radical concepts like false consciousness. For instance, Maeve Cooke begins a

recent defense of the idea of “objectively necessary false consciousness” by affirming the “ethical intuition that the freedom of human beings consists in important measure in the freedom to pursue their conceptions of the good on the basis of reasons that they are able to call their own.”<sup>3637</sup> Regardless of its ideological origins, this ethical intuition is now close to non-negotiable among those concerned to vindicate the concept of false consciousness. In what follows, then, I will assume that any successful conception of deep false consciousness must start with an account of real interests that is compatible with the idea of democratic toleration outline above, but I will not provide any new arguments in defense of the idea.

Now, the common but misguided argument against the idea of deep false consciousness that I want to discuss in this section involves denying that there exists a substantive account of real interests capable of satisfying the defensibility requirement. Stephen White, insisting on respect for the “fact of pluralism,” offers a version of this objection:

[a]ny social theory today that builds in from the start a specific substantive conception of the real interests of any large category of actors is simply going to be too restrictive in terms of both what those actors might constitute as problems and what they might choose as legitimate resolutions of those problems.<sup>38</sup>

For White, the assignment of real interests to the targets of one’s criticism always involves an arbitrary choice of the sort Lukes complained about in the example of the American electorate. On the other side of the debate, defenders have responded to this sort of objection by arguing that while some accounts of real interests, including accounts historically associated with theories of false consciousness, may indeed be arbitrarily restrictive, it is possible to offer an account of interests that satisfies the defensibility requirement and, thus, salvages the idea of deep false consciousness. As I’ll argue below, I think the defenders get the better of this particular debate, but in so doing reveal some inescapable limitations of the idea of deep false consciousness.

In order to understand the thinking of both critics and defenders of the idea of deep false consciousness, it will be helpful to start with a sketch of the “orthodox” conception of false consciousness to which both groups are, in different ways, responding. This orthodox conception has its roots in Marx’s thought. Though Marx never used the term “false consciousness” (the coinage is usually credited to Engels who used the term in a letter to Franz Mehring in 1893), he evidently endorsed the underlying tenets of the idea. Consider the following passage from “The Holy Family” in which Marx speaks to the aims of the proletariat:

The question is not what this or that proletarian, or even the whole of the proletariat at the moment considers as its aim. The question is what the proletariat is, and what, consequent on that being, it will be compelled to do. Its aim and historical action is irrevocably and obviously demonstrated in its own life situation as well as in the whole organization of bourgeois society today.<sup>39</sup>

This passage, often cited as evidence for the claim that Marx held at least an implicit notion of what would eventually come to be called false consciousness, is notable in a couple of respects.<sup>40</sup> First, we see the suggestion of a particular standard against which the subjective aims of workers are to be measured, namely, the objective aims of workers as a class. Second, we see the suggestion of a particular method for determining that standard, namely, a theoretically informed analysis of society that is supposedly capable of revealing the historical mission of the proletariat. Marx is clear that the subjective aims of workers, considered individually or collectively, carry no weight when it comes to determining the objective aims of the class. What is needed is not input from workers themselves but a correct social theory.<sup>41</sup>

These ideas would be worked into an explicit conception of false consciousness by Georg Lukács, who theorized false consciousness in conjunction with what he called the “problem of organization,” which is, roughly, the problem of bringing workers’ subjective beliefs and attitudes into alignment with the beliefs and attitudes appropriate to their class position. Lukács claims that

by relating consciousness to the whole of society it becomes possible to infer the thoughts and feelings which men would have in a particular situation if they were able to assess both it and the interests arising from it in their impact on immediate action and on the whole structure of society. That is to say, it would be possible to infer the thoughts and feelings appropriate to their objective situation...Now class consciousness consists in the fact of the appropriate and rational reactions imputed to a particular typical position in the process of production...This analysis establishes right from the start the distance that separates class consciousness from the empirically given, and from the psychologically describable and explicable ideas, which men form about their situation in life.<sup>42</sup>

Lukács shares with Marx the view that a theoretically informed view of class conflict enables one to identify the objective interests of those who are party to the conflict. As he suggests, “the naïve description of what men in fact thought, felt and wanted at any moment in history” is “merely the material of genuine historical analysis.”<sup>43</sup> And it is this analysis, not the desires or beliefs of workers themselves, that will reveal the real interests of the proletariat. But Lukács goes further than Marx in drawing out the implications of this view. He declares the possibility of determining “the thoughts and feelings appropriate to [workers'] objective situation,” using the term “class consciousness” to label the complex of thoughts, feelings, and attitudes that workers ought to hold by virtue of their class position. For Lukács, “class consciousness” supplies the standard of real interests necessary to judge the “empirically given” consciousness of workers. To the extent that workers’ “subjectively justified” thoughts and feeling diverge from this standard, they constitute what Lukács regards as objectively “false consciousness.”<sup>44</sup>

Now, there are some obvious difficulties with the standard of real interests that underwrites this conception of false consciousness—difficulties both with the substance of the standard and the way in which it is derived. With respect to the substance of the standard, the equation of real interests with class interests tends to strike contemporary audiences as arbitrarily restrictive. One might expect committed liberals to object to the equation of real interests with class interests, but the equation looks problematic even from the perspective of those committed to a more solidaristic politics. As Seyla Benhabib notes, “[s]truggling collectivities may be

formed around other normative concerns besides class interests, and relations of exploitation and domination may be based upon other characteristics like sex, race, ethnic and linguistic identity, and even age.”<sup>45</sup> (As noted above, these expressions of respect for pluralism have become standard for theorists on the left.) The problem with the orthodox conception of false consciousness, Benhabib suggests, is that it denies the legitimacy of these other forms of concern and identification. Benhabib’s observations underline the significance of the fact, noted above, that persons have multiple and conflicting interests by virtue of the fact that they have complex identities. To insist that real interests are necessarily class interests is to deny—or, at the very least, to significantly downgrade—the normative significance of all but one of the many elements that form an individual identity.

With respect to the way in which the orthodox conception’s standard of interests is derived, it appears that a theorist like Lukács’s claim to speak for the real interests of workers reflects little more than an “arrogant assumption of superior knowledge” on his part.<sup>46</sup> Of course, there is nothing inherently “arrogant” about assuming superior knowledge in some domain. We do not typically question the doctor’s assumption of superior knowledge when she tells us that a certain course of treatment is in our best interest. The problem with the theorist of false consciousness’s assumption of superior knowledge is that it appears unwarranted. Even if one accepts that the right social theory would allow the theorist to determine society’s present phase of development in relation to some discoverable endpoint, it is not clear how this knowledge would entitle him to speak on behalf of the real interests of people living here and now. At best, the right social theory would provide one with the ability to say something about the interests that a group of, say, workers have *qua* workers given the direction in which the world is heading, but this fails to get beyond the problem of complex identities discussed in the previous

paragraph. Being able to determine the “thoughts and feelings appropriate” to a group of workers would require one to know not just what is good for workers in general, but what is good for a particular, historically-situated group of people who happen to be workers but whose identities are multifaceted and who may be invested in a variety of more or less valuable projects. It is this second sort of implicit knowledge claim that the theorist of false consciousness, even one armed with a compelling social theory, has trouble justifying.<sup>47</sup> More troubling is the fact that the theorist of false consciousness’s claim to know the interests of others seems to be offered in a way that resists falsification, thereby insulating the theorist from criticism. Recall Lukács’s claim that what “men in fact thought, felt and wanted” can provide no counterevidence against the results of the theorist’s “genuine historical analysis.” This approach has been called “epistemologically authoritarian” in the sense that “what counts as a rational interest to the epistemically privileged theorist...cannot be challenged by the human subjects concerned.”<sup>48</sup>

As I suggested above, these difficulties with the orthodox conception provide the starting point for both critics and defenders of the idea of deep false consciousness. Both camps agree that the orthodox conception is flawed, but split on the question of whether its flaws are inherent to the very idea of deep false consciousness or unique to a certain articulation of that idea. There are those, like Colin Hay, who assume that any conception of deep false consciousness will reproduce the flaws of the orthodox conception. Hay suggests that the very idea of deep false consciousness “conjures up” a

deeply condescending conception of the social subject as an ideological dupe...Not only is this wretched individual incapable of perceiving her/his true interests...But rising above the ideological mists which tame the masses is the enlightened academic who from his/her high perch in the ivory tower may look down to discern the genuine interests of those not similarly blessed. Clearly such a conception is both logically unsustainable and politically offensive.<sup>49</sup>

Though his venom is notable, the substance of Hay’s charge that the idea of false consciousness is “politically offensive” and “logically unsustainable” is more or less in line with the criticisms



analyzed above: the orthodox conception is “politically offensive” in the sense that its standard of real interests implies that there is only one form of legitimate political identification; it is “[epistomo-]logically unsustainable” in the sense that its derivation of that standard depends on an unwarranted assumption of superior knowledge. If one assumes that any conception of false consciousness must traffic in these sorts of illiberal claims and epistemically dubious methods, it’s easy enough to see why he would be inclined to dismiss the very idea as fundamentally flawed. However, the assumption that attributes the flaws of the orthodox conception to the broader idea of deep false consciousness is clearly problematic. Deep false consciousness must be defined in opposition to some substantive standard of real interests, as we’ve seen. But there are as many possible conceptions of deep false consciousness as there are accounts of real interests. There is no obvious reason to assume that because some accounts are indefensible, all of them must be.<sup>50</sup>

Of course, even if the idea of deep false consciousness cannot be dismissed on the basis of the easy, uncritical assumption that the imputation of interests to others is never defensible, it is left to those who want to salvage the idea to show what a defensible account of real interests would look like. Defenders have responded to this challenge by attempting to re-conceptualize the idea of deep false consciousness atop thinner accounts of real interests, proceeding from the intuition that a thinner account of interests will place fewer restrictions on what counts as a legitimate set of aims, thus avoiding the charges of arbitrariness invited by the orthodox conception. Thinning down one’s account of real interests also apparently addresses some of the worries about epistemic foundations that dog the orthodox conception, since a thinner account of real interests is more likely to be justifiable without resort to implausible claims of epistemic privilege. Denise Meyerson, who has written one of the more thoroughgoing defenses of the idea

of deep false consciousness, is typical in this regard. She argues that we ought to look not to arcane social theories, but to psychology for “knowledge of the facts relevant to interests.”<sup>51</sup> As she puts it, “psychology may tell you that it will only bring you distress if you persist in trying to follow in an ambitious father’s footsteps; or in obsessively pursuing wealth; or if you are committed to a life of duplicity with its attendant costs of hypocrisy and lack of spontaneity.”<sup>52</sup> Meyerson readily admits that there is some controversy about “what psychological theory is true.” But her argument proceeds at a level of abstraction from which the question of which psychological theory is true can apparently be set aside. After all, she notes, “those who deny that there are facts about interests do not usually rest their case on the claim that psychology is unable to yield objective truths. Usually they accept that psychology can provide factual truths and think merely to deny that anything evaluative could follow from them.”<sup>53</sup> Meyerson’s response is straightforward: if a true theory of psychology can show what “causes [a] person to function effectively,” then a true theory of psychology will reveal facts about what is in a person’s interests.

In place of “class consciousness,” then, Meyerson offers an account of real interests grounded in our shared “psychological nature.”<sup>54</sup> The move seems plausible enough if the goal is simply to refute sweeping objections to the idea of deep false consciousness by showing that there are at least *some* accounts of objective interests that can be defensibly imputed to individuals who do not recognize those interests as their own. Meyerson spends little time speculating about which particular interests arise from our shared psychological nature (a telling decision I’ll revisit below) but judging from the examples she provides, the sorts of interests she has in mind are indeed less restrictive than the interests associated with class consciousness and, thus, less likely to appear “politically offensive.” Moreover, Meyerson’s claim that even critics

of the idea of false consciousness tend to accept that “psychology can provide factual truths” about what causes a person to function effectively serves to show that imputations of interests can be defended with reference to relatively established and defeasible canons of knowledge and need not depend on the kind of “epistemic authoritarianism” associated with the orthodox conception.

Meyerson’s cautious, psychologically grounded conception of false consciousness thus shows the error of simply assuming that any imputation of real interests must involve an unsustainable claim of epistemic privilege. However, what I want to suggest is that Meyerson comes well short of vindicating the idea of deep false consciousness, and in fact ends up revealing some of its fundamental limitations. To see why this is the case, we need to step back for a moment in order to consider what a truly successful conception of false consciousness would look like. If defensibility were a sufficient condition for a successful conception, Meyerson would arguably succeed, at least so long as one accepts that a true psychological theory can reveal facts about what it means for a person to function effectively. But defensibility must be regarded as a necessary rather than a sufficient condition for a successful conception of deep false consciousness. In addition to being underwritten by a defensible account of real interests, a successful conception of deep false consciousness must furnish some determinate standards for criticizing individuals’ politically relevant beliefs and attitudes. This requirement—call it the determinacy requirement—is tied to the purpose of false consciousness as a critical concept. As I noted in the introduction, conceptions of false consciousness are diagnostic tools: they serve to identify certain politically relevant beliefs and attitudes as flawed—i.e., as constituting a false consciousness. If the interests imputed to various groups of people by a particular conception of false consciousness are indeterminate with respect to relevant political

and social questions, the conception will have little value as a critical tool, even if its underlying account of interests is defensible.

The problem with Meyerson's conception of false consciousness is that it is hard to see how the interests grounded in our shared psychological nature tell us anything about what sort of political beliefs and attitudes we ought to hold. Perhaps there are certain forms of political identification that are so ill-suited to our psychological nature that adopting them would frustrate our interests, but that seems doubtful given what little Meyerson has to say about the nature of our psychological interests. Now Meyerson herself, like other defenders of deep false consciousness, does not see this as a problem. This is not particularly surprising. As I've tried to show, the troubled legacy of the orthodox conception has helped to set a particular agenda for contemporary debates about false consciousness. Since the orthodox conception's most glaring shortcoming is that it begins with an unsustainable standard of real interests, defenders of the idea have focused their efforts on shoring up this weakness by striving to develop accounts of real interests that the theorist of false consciousness need not be embarrassed to endorse. One could say in Meyerson's defense, then, that determinacy is just not a relevant consideration for her. But this argument is unconvincing. Defenders of the idea of deep false consciousness can always outflank critics by developing thinner and thinner accounts of real interests until they light upon one which can be plausibly defended. But if the accounts become so thin that they are no longer capable of providing critical leverage for evaluating individuals' politically relevant beliefs and desires then the idea of false consciousness has been emptied of its critical value. Put differently, if the only way to respond to the charge that the idea of deep false consciousness is "politically offensive" is to make the idea politically irrelevant, then defenders of the idea have lost the debate.

Of course, even if I'm correct that a conception of deep false consciousness grounded in our shared psychological nature is too insubstantial to have much critical value, one might respond that this is a fault of Meyerson's particular conception of false consciousness rather than a limitation of the idea itself. But I think Meyerson's neglect of the determinacy requirement is less an oversight on her part than a product of a fundamental tension between the defensibility and determinacy requirements. The tension exists because, as Meyerson and others understand, in order to satisfy the defensibility requirement a conception of false consciousness must be built upon a relatively thin account of interests. On the other hand, if an account of imputable interests is to furnish a set of determinate standards for criticizing agents' actual beliefs and desires it will need to be relatively thick. In short, the defensibility and determinacy requirements pull in opposite directions, creating a dilemma for the theorist of false consciousness. It is not simply that Meyerson satisfies the defensibility requirement while neglecting the determinacy requirement; she satisfies the defensibility requirement by neglecting the determinacy requirement. Understanding the nature of this dilemma helps to put the orthodox conception, rejected by both defenders and critics of deep false consciousness as obviously wrongheaded, in the proper perspective. Thinkers like Marx and Lukács simply resolve the dilemma in favor of the determinacy requirement by offering an account of real interests that provides determinate standards for political and social criticism, but that is indefensible in the estimation of most contemporary thinkers. Of course, there is no dilemma from Marx and Lukacs's perspective, since neither is concerned to build a theory that accommodates the fact of pluralism or acknowledges the value of democratic toleration. But the dilemma is inescapable for those concerned to render talk of real interests consistent with democratic values. By re-conceptualizing the idea of deep false consciousness atop thinner accounts of real interests, latter

day defenders of the idea do not escape the dilemma of real interests; they simply move from one horn of the dilemma to the other.

Ultimately, in evaluating the idea of deep false consciousness, the question we should ask is not simply “are outside observers ever justified in imputing interests to the members of different groups?” Though some critics try to argue against the idea of deep false consciousness by answering “no” to this question, this is not the place to take a stand against the idea. As Meyerson’s argument helps to bring out, it is implausible to say that there is nothing that is objectively good for persons that can be known to outside observers. The question we should be asking is more complex: if outside observers are justified in imputing interests to the members of various groups, will those interests be substantial enough to furnish a set of determinate standards for criticizing those persons’ politically relevant beliefs and attitudes? My answer to that question is largely negative: any conception of deep false consciousness that satisfies the defensibility requirement interpreted in light of democratic ideals will fail fully to satisfy the determinacy requirement. I emphasize the qualifier “fully” in the previous sentence, since it is possible, I think, that a conception of deep false consciousness might be defensible and at least somewhat determinate with respect to salient issue of political and social organization. There may be cases of what White calls “extreme domination” in which even a thin account of real interests provides enough leverage to criticize the behavior of affected individuals, but one imagines that such cases would be relatively rare in modern democracies.<sup>55</sup>

## **6. Shallow False Consciousness**

I turn now to the idea of shallow false consciousness, evaluating it along the same dimensions that I evaluated deep false consciousness, first asking what it takes to justify attributions of

shallow false consciousness, and then asking whether the idea furnishes a useful tool for critical social inquiry.

As noted above, attributions of shallow false consciousness raise a different kind of justificatory burden than attributions of deep false consciousness. Where the latter require a defensible account of real interests generated by some theory of human flourishing, the former require a defensible means of revealing the gap between one's underlying aims and his more immediate preferences. Thus, while the idea of deep false consciousness raises difficult philosophical questions about the nature of the good life, the idea of shallow false consciousness raises a difficult empirical question: how can we justify claims about a divergence between one's preferences and his underlying aims when we lack the ability to observe those aims directly? I doubt that there is a single best answer to this question. Rather, there are different methods and different types of evidence that might be used to generate inferences about citizens' underlying aims. Whether or not such inferences are plausible in any particular instance will depend on the strength of the evidence at hand and the manner in which it is interpreted. For this reason, generalizations about the justifiability of attributions of shallow false consciousness are difficult to make. Out of necessity, then, my approach here will be rather modest. Looking across some recent works of empirical political science, I want to consider a few possible approaches that have been used—and might be profitably used in the future—to judge the extent to which individuals' expressed preferences diverge from their underlying aims. The three approaches I discuss can be distinguished in terms of the type of evidence they seek to provide: testimonial, experimental, and statistical.

[The first and perhaps most obvious way of assessing the fit between persons' underlying aims and their preferences is by asking them why they hold the preferences they do. Though the

practice of asking people to share their reasons for holding certain views is more typical of journalistic treatments of public opinion than it is scholarly studies of the subject, political scientists have begun to adopt the practice. For instance, Katherine Walsh begins her recent study of rural voters in Wisconsin with the question at the very heart of debates about false consciousness and democracy—“Why do people vote against their interests?”—and sets about answering that question by conducting long form interviews with (and observing discussions among) small groups of Wisconsin residents.<sup>56</sup> Though Walsh’s project is descriptive rather than critical in the sense that she remains agnostic about the coherence of the reasons her participants offer for their beliefs and preferences, her participants’ general opposition to redistributive measures is premised on a number of factual claims (e.g., about where tax dollars are spent) that might be subjected to scrutiny. There are, of course, some limitations to this approach. The time intensiveness of interviews eliciting not just preferences but the reasons behind those preferences will generally preclude investigators from recruiting a large number of participants into a single study. Thus, in a study like Walsh’s, where a limited number of participants are drawn from a relatively small geographic area, results are difficult to generalize. Additionally, there is some concern that the rationales offered by interview subjects will misrepresent the actual structure of their beliefs, though the extent to which this is a problem will depend heavily on the construction of the interview protocol and the manner in which discussions are allowed to proceed.

A second approach to the problem involves turning to experimental evidence. In studies of this sort, investigators begin by asking participants to give their untutored preferences over some range of political issues. Then, after exposing participants in a treatment group to factual information germane to those issues, investigators ask participants to provide their preferences a second time. Variations on this approach are used, for instance, by Kuklinski et al. and Mettler



and Guardino in their analyses of Americans' attitudes towards welfare spending, and by Gilens in his evaluation of "preferences for spending on prisons, foreign aid, and the environment."<sup>57</sup> The guiding assumption of such studies is that to the extent that preferences become better informed, they become more representative of underlying interests. In principle, on this way of thinking, fully informed preferences—that is, the preferences persons would have if they knew all relevant information and had the time and capacity to process that information—would be perfectly representative of underlying interests. By measuring preference shifts either "between subjects" (in the case of comparisons between treatment and control groups) or "within subjects" (in the case of comparisons between the pre- and post-treatment preferences of members of the same group), researchers are able to generate inferences about the extent to which untutored preferences are refracted through ignorance and error. Of course, studies of this sort have their limitations as well. While it is fairly easy to tell *when* exposure to new information affects preferences, it is sometimes difficult to tell *how* information affects preferences. When preferences shift as a result of exposure to new information, one interpretation is that participants have used the information to reformulate their preferences to bring them into closer alignment with underlying interests. But it is possible that shifts in preference are actually the result of "a momentary response to the experimental stimulus." As Gilens explains, "this may occur if exposure to the information influences responses by 'priming' one aspect of the issue rather than by prompting a genuine reevaluation of the issue in light of new information."<sup>58</sup> This issue can be mitigated by ensuring that the control group is exposed to similar priming effects, but it is difficult to overcome completely.

In addition to testimonial and experimental evidence, researchers have used statistical evidence to generate inferences about how well citizens' preferences line up with their

underlying interests. One approach involves using survey data that includes measures of general political and/or issue specific knowledge to compare the preferences of well-informed members of a defined demographic group to the preferences of poorly-informed members of that same group. Assuming that all members of the group share a similar set of core commitments, divergence in their observed preferences is taken to reflect the fact that poorly informed group members struggle to identify the sorts of policies that would advance their underlying interests. To be sure, these sorts of intra-group comparisons come with some concerns. In particular, the idea that well-informed and poorly-informed members of a defined demographic group differ only in terms of their knowledge of salient political facts is implausible. As Gilens puts it, “respondents who correctly answer policy-specific questions differ in a host of ways from those who do not. And despite the long list of control variables used in these analyses, we cannot hope to identify and adequately measure all the characteristics that might distinguish these two groups.”<sup>59</sup> However, not every statistical approach to the problem involves these sorts of intra-group comparisons. Bartels uses a statistical model to show that the votes of low income voters are substantially influenced by the economic fortunes of the wealthiest Americans and election year economic growth. Bartels chalks this up to a form of “false consciousness,” in which voters intending to vote on the basis of their economic interests respond to the wrong economic indicators and end up casting votes which frustrate those interests. Of course, to call this a form of false consciousness and a failure of economic accountability, as Bartels does, is to assume that low income voters do not simply intend to vote for candidates whose economic policies lead to election year economic growth and high-income growth. Though Bartels’ assumption seems more plausible than the alternative story, it bears emphasizing that it represents a particular, defeasible interpretation of the available data.]

Because subsequent chapters of the dissertation will examine in far greater detail the application of these methods to specific case studies, I will, for now, offer just a few general observations on the methods discussed here and the prospects for arriving at justified attributions of shallow false consciousness. First, because no one method of reaching inferences about how well persons' preferences match their underlying interests is without faults, attributions of shallow false consciousness are more credible to the extent that they can be substantiated by more than one method. While there may be reason to doubt the results of any one study, when multiple studies employing diverse methods converge on similar findings, we have good reason to take those findings seriously. Secondly, if we view the problem through the lens of democratic theory and institutional design, the ability to say conclusively that any one person's preferences are distorted by ignorance or error is less important than the ability to identify patterns and tendencies across a wider population, since any institutional solution to the problem will not be tailored to individual cases but designed to address broad causes. Lastly, whatever their particular limitations, these methods for justifying claims about shallow false consciousness are not epistemically authoritarian in the sense described above—that is, they need not be presented in a way that shields their conclusions from scrutiny by the groups in question or by outside observers. On the contrary, as I've already briefly shown, key assumptions can be rendered transparent and competing interpretations of relevant data can be offered and debated.

Assuming for the moment, then, that it will sometimes be possible to justify attributions of shallow false consciousness, there remains a question about the critical value of those attributions. After all, I've said that efforts to re-conceptualize deep false consciousness atop thinner accounts of real interests empty the idea of its critical force. A similar objection might be put to the idea of shallow false consciousness. One might grant that while shallow false

consciousness is superficially an idea of false consciousness—in the sense that it allows for the criticism of beliefs and desires based on their incompatibility with one’s “interests”—the critical force of the concept is diminished in the switch from an evaluative criterion based on objective interests to one based on subjective interests. False consciousness was developed into a critical concept by theorists dissatisfied with what has variously been called the “liberal,” “positivist,” and “old behaviorist” practice of equating interests with what people “actually want and prefer.”<sup>60</sup> For theorists of false consciousness, the problem with this equation is that it fails to take seriously the many ways in which people’s wants may be shaped by ignorance, irrationality, and manipulation.<sup>61</sup> Because subjective interests are themselves grounded in wants, the idea of shallow false consciousness may look like a retreat to a paradigm whose limitations the concept of false consciousness was meant to overcome. Furthermore, because the idea of shallow false consciousness targets beliefs and desires generated by errors in instrumental reasoning, it provides no leverage for criticizing beliefs and desires that faithfully express underlying values, even seemingly problematic values. This aspect of the idea may appear to generate some puzzling conclusions. For instance, the wage worker whose political preferences are an undistorted expression of his free market values—a clear victim of false consciousness on the orthodox view—could not be said to suffer from what I’m calling shallow false consciousness. Taken together, these concerns may suggest that even if one can make defensible claims about gaps between someone’s subjective interests and some subset of his beliefs and preferences, those claims are insufficient to ground a distinct idea of false consciousness.

I’ll address both parts of this objection in turn, starting with the thought that an idea of false consciousness must be grounded in a “trans-subjective evaluative criterion” in order to maintain its distinctive critical perspective, before concluding with some forward looking

thoughts on the value of the idea. While it's true that the idea of shallow false consciousness retreats from the search for an evaluative criterion based on objective interests, it does not retreat to the "uncritical" position of equating interests with expressed preferences. As we've already seen, subjective interests indeed are grounded in wants, but we can disaggregate the broad category of "wants" into relatively deep-rooted aims and higher order desires and preferences that are instrumentally related to those aims. It's possible, then, to talk about preferences and subjective interests coming apart in much the same way that we talk about preferences and objective interests coming apart. Though the idea of shallow false consciousness rejects the sort of essentialism that complicates the idea of deep false consciousness, it also rejects "the old behaviorist commitment" of taking expressed preferences at face value. It occupies a critical standpoint between these two positions, a standpoint from which it is possible to diagnose preferences distorted by various forms of misapprehension but not possible to say what aims individuals ought ultimately to pursue. Even from this middling standpoint, however, the idea of shallow false consciousness fulfills the critical function of identifying distortions in the self-understanding of social actors.

But what about the idea that it is a shortcoming of shallow false consciousness that the idea can't be used to criticize beliefs and desires that faithfully express problematic underlying values? The first thing to note in response to this sort of concern is that the specification of a new critical category does require the removal of any other critical categories from our vocabulary. Beliefs and desires that cannot be criticized as expressions of shallow false consciousness can still be criticized for being immoral, selfish, imprudent, and so forth. Even the idea of deep false consciousness remains available to those who would criticize, from the perspective afforded by some account of real interests, the ends persons set for themselves. I've already explained why I

think the scope for such criticism is limited, but the specification of shallow false consciousness as a distinct critical category does nothing to further those limitations. As I suggested at the outset, my aim is to replace the bloated concept of false consciousness with distinct “ideas that we can choose to use or not, depending on how well the use of these ideas suits our investigative purpose.” That no one of those ideas captures the entire range of phenomena that have been lumped together under the heading of false consciousness is not an unfortunate consequence, but rather the goal of this exercise.

Ultimately, the question of whether the idea of shallow false consciousness offers a useful critical tool can be answered only by seeing how well it “suits our investigative purposes” in some concrete cases. That, however, remains a task for another paper.

---

<sup>1</sup> Elizabeth Anderson, “Unstrapping the Straightjacket of ‘Preference’: A Comment on Amartya Sen’s Contributions to Philosophy and Economics,” *Economics and Philosophy*, no. 17 (2001): 21.

<sup>2</sup> *Ibid.*, 22.

<sup>3</sup> Kevin S. Decker, “Perspectives and Ideologies: A Pragmatic Use for Recognition Theory,” *Philosophy & Social Criticism* 38, no. 2 (February 1, 2012): 215–26; Michael Rosen, “On Voluntary Servitude and the Theory of Ideology,” *Constellations* 7, no. 3 (September 1, 2000): 393–407; Steven Lukes, “In Defense of ‘False Consciousness,’” *University of Chicago Legal Forum*, 2010; Raymond Geuss, *The Idea of a Critical Theory: Habermas and the Frankfurt School* (Cambridge University Press, 1981); Frank Cunningham, *Democratic Theory and Socialism* (Cambridge University Press, 1987).

<sup>4</sup> Additionally, the term “consciousness” implies that the cognitive and motivational errors in question display a certain comprehensiveness and cohesiveness. In other words, “false consciousness” suggests more than a few isolated mistakes; it suggests a way of seeing the world, or some part of it, that is shot through with error. That said, I do not think it is possible to draw a sharp distinction between false consciousness proper and some less comprehensive combination of cognitive and motivational error; the two are different in scope rather than kind.

<sup>5</sup> These concerns are central to the work of Steven Lukes, who has been one of the more resolute defenders of the concept of false consciousness in recent years. Lukes’ well-known analysis of power aims to reveal the ways in

---

which power can operate behind the scenes, as it were, to shape the desires of subordinated populations. As he puts it, “the supreme exercise of power [is] to get another or others to have the desires you want them to have.” See Steven Lukes, *Power: A Radical View*, 2nd ed. (Palgrave Macmillan, 2005), 27. In the context of this analysis, Lukes associates “false consciousness” with a “cognitive power of considerable significance and scope: namely, the power to mislead” (149). Though he doesn’t put the point quite this way, what Lukes calls the “third face of power” is essentially the power to induce false consciousness in others—that is, the power to induce others to hold desires that are contrary to their own interests. While Lukes is particularly interested in the ways in which powerful agents may be able to induce false consciousness in others, it is no part of the definition of false consciousness that the condition must have an external cause. Lukes himself acknowledges that false consciousness may have internal causes as well—a possibility that has been extensively studied by social psychologists like John Jost, who argues that false consciousness may arise spontaneously in subordinated populations as a means of managing cognitive dissonance. See John T. Jost, “Negative Illusions: Conceptual Clarification and Psychological Evidence Concerning False Consciousness,” *Political Psychology* 16, no. 2 (June 1995): 397.

<sup>6</sup> See, e.g., Ronald Dworkin, Mark Lilla, David Bromwich, and Jonathan Raban, “The Historic Election: Four Views,” *The New York Review of Books*, December 9, 2010, <http://www.nybooks.com/articles/archives/2010/dec/09/historic-election-four-views/>; Niall Ferguson, “Viewpoint: Why the Young Should Welcome Austerity,” *BBC*, June 17, 2012, sec. World, <http://www.bbc.co.uk/news/world-18456131>; Jeffrey A. Winters and Benjamin I. Page, “Oligarchy in the United States?,” *Perspectives on Politics* 7, no. 04 (2009): 731–51; Susanne Lohmann, “An Information Rationale for the Power of Special Interests,” *The American Political Science Review* 92, no. 4 (December 1, 1998): 809–27; Michael X. Delli Carpini and Scott Keeter, *What Americans Know about Politics and Why It Matters* (Yale University Press, 1997).

<sup>7</sup> Of course, to answer adequately the question of what false consciousness means for democracy, we would need to develop a list of democratic desiderata and show how these desiderata are impacted by the phenomenon under consideration. This task—which I take up in earnest in chapter four of the dissertation—is complicated by the fact that different conceptions of democracy vary widely with respect to the competencies they demand of democratic citizens and the outcomes they expect from the democratic process. For now, I proceed under the assumption that whatever system of democracy one favors, the existence of something like false consciousness poses a plausible threat to the legitimacy of the outcomes generated by that system. Though this assumption motivates this chapter and the larger project of which it is a part, nothing I say here hangs on its being true.

<sup>8</sup> lukes

<sup>9</sup> For some, the concept’s historical association with totalitarian political movements is reason enough to hold it in disrepute. For others, particularly followers of Michele Foucault, the idea that all our beliefs and values are the products of socially constructed regimes of truth renders talk of true and false consciousness essentially meaningless. Still others challenge the explanatory adequacy of theories of false consciousness, arguing that behavior which is said to be the product of false consciousness can usually be shown to have other causes.

<sup>10</sup> Clarissa Hayward and Steven Lukes, “Nobody to Shoot? Power, Structure, and Agency: A Dialogue,” *Journal of Power* 1, no. 1 (2008): 5–20.

<sup>11</sup> 397

<sup>12</sup> Richard E. Flathman, *The Public Interest: An Essay Concerning the Normative Discourse of Politics* (New York: John Wiley & Sons, 1966), 15.

<sup>13</sup> Jane J Mansbridge, *Beyond Adversary Democracy* (Chicago: University of Chicago Press, 1983), 24.

<sup>14</sup> Hanna Pitkin, *The Concept of Representation* (University of California Press, 1967), 156.

<sup>15</sup> These are adapted from the OED. See “Interest,” *OED Online* (Oxford University Press), accessed January 11, 2014, <http://www.oed.com/view/Entry/97735>.

<sup>16</sup> Isaac Balbus, “The Concept of Interest in Pluralist and Marxian Analysis,” *Politics and Society* 151, no. 1 (1971): 152.

<sup>17</sup> This example is adapted from Balbus (1971).

<sup>18</sup> William E. Connolly, “On ‘Interests’ in Politics,” *Politics & Society* 2, no. 4 (September 1, 1972): 471.

<sup>19</sup> Ibid.

<sup>20</sup> Connolly, “On ‘Interests’ in Politics.”

<sup>21</sup> Joel Feinberg, *Harm to Others* (Oxford University Press, 1984).

<sup>22</sup> Ibid., 40.

<sup>23</sup> In democratic politics, citizens often face similar tradeoffs. As Roemer points out, citizens do not typically vote on discrete policies; rather, they vote on “policy bundles” offered by competing parties and their candidates. When selecting between policy bundles, citizens may find themselves in the position of Feinberg’s professor, forced to

---

advance one interest at the expense of the other. This is one reason why it is often difficult to show that it is clearly in A's best interest to vote for Candidate X.

<sup>24</sup> Denise Meyerson, *False Consciousness* (Oxford University Press, USA, 1991), 124.

<sup>25</sup> Note here about Pitkin's objections to purely objective interests..

<sup>26</sup> Feinberg, *Harm to Others*, 45.

<sup>27</sup> This is roughly the way Brian Barry uses the term "interest" when he claims that "to say...that an action or policy is in somebody's interests is not actually to say that it satisfies his immediate wants at all; it is rather to say that it puts him in a position to satisfy his wants." To return to the example of my striving to become a successful politician, what is "in my interests" are those actions which will help me become a politician. Those actions are in my interests *objectively* whether or not I am subjectively motivated to pursue them. And yet, they are in my interest objectively only because I have certain subjective goals and aspirations. If I had different goals, a different set of things would be in my objective interests. Brian Barry, *Political Argument: A Reissue With a New Introduction* (University of California Press, 1965), 183.

<sup>28</sup> Robert Paul Wolff, "Methodological Individualism and Marx: Some Remarks on Jon Elster, Game Theory, and Other Things," *Canadian Journal of Philosophy* 20, no. 4 (December 1, 1990): 480.

<sup>29</sup> Ibid.

<sup>30</sup> Kelly point here about how some critics of false consciousness assume that it must imply utter irrationality. Perhaps heath as well.

<sup>31</sup> Meyerson, *False Consciousness*.

<sup>32</sup> One additional point worth noting is that second order error can occur at different levels of reasoning. Some who fails to see how to advance a subjective interest in economic sufficiency might suffer from fundamental confusion about how markets work, how wealth is generated and sustained, and so forth. Or he may possess this sort of foundational knowledge and but make simple and specific mistakes in selecting a strategy for achieving his underlying goal. In both cases, some type of cognitive error prevents him from seeing how best to advance his subjective interest—which is why both cases amount to instances of second order error—but the nature of the error is quite different in the two cases. Second order error is thus a fairly gross category that encompasses multiple sorts of error.

<sup>33</sup> Joseph Raz, *The Morality of Freedom* (Oxford University Press, 1986), 143.

<sup>34</sup> Arguably, the distinction between deep and shallow false consciousness has a structure similar to Williams' distinction between internal and external reasons. Shallow false consciousness might be construed as a failure to follow one's internal reasons, whereas deep false consciousness might be construed as a failure to follow external reasons. Though in the case of deep false consciousness, the external reasons would not be moral reasons but reasons rooted in one's own good. See: Bernard Williams, "Internal and External Reasons," in *Rational Action*, ed. Ross Harrison (Cambridge University Press, 1979), 101–13.

<sup>35</sup> Lukes, "In Defense of 'False Consciousness.'"

<sup>36</sup> Maeve Cooke, "Resurrecting the Rationality of Ideology Critique: Reflections on Laclau on Ideology," *Constellations* 13, no. 1 (March 1, 2006): 4.

<sup>37</sup> See also decker.

<sup>38</sup> Stephen K. White, "The Very Idea of a Critical Social Science," in *The Cambridge Companion to Critical Theory*, ed. Fred Rush (Cambridge University Press, 2004), 317.

<sup>39</sup> Karl Marx, "The Holy Family," in *Karl Marx: Selected Writings*, ed. David McLellan, 2nd ed. (Oxford University Press, USA, 2000), 149.

<sup>40</sup> Meyerson, *False Consciousness*; Cunningham, *Democratic Theory and Socialism*.

<sup>41</sup> For an objection to the idea of basing claims about human well-being in "the correct social theory" and the persistence of the practice in Marxist theory, see: Bernard E. Harcourt, *Radical Thought from Marx, Nietzsche, and Freud, Through Foucault, to the Present: Comments on Steven Lukes' "In Defense of False Consciousness,"* SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, June 16, 2011), 11, <http://papers.ssrn.com/abstract=1865962>.

<sup>42</sup> György Lukács, *History and Class Consciousness: Studies in Marxist Dialectics* (MIT Press, 1971), 51.

<sup>43</sup> Ibid.

<sup>44</sup> Ibid., 50.

<sup>45</sup> Seyla Benhabib, *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory* (New York: Columbia University Press, 1986), 129.

<sup>46</sup> Lukes, "In Defense of 'False Consciousness.'"



---

<sup>47</sup> As S.I. Benn observed: “When we talk about the interests of a class, as opposed to an organized group, we construct a paradigm, what anyone might be expected to want, or to be concerned about, insofar as he was simply a farmer (and not for instance a member of the Union executive committee, a vegetarian or a prosperous philosopher with a rural hobby, any of which uncharacteristic features might be associated with interests which conflict with the paradigm.” I would simply add that the problem Benn identifies is not peculiar to a few prosperous philosophers but quite widespread. Put differently, we rarely encounter the individual who is “simply a farmer.”

<sup>48</sup> Cooke, “Resurrecting the Rationality of Ideology Critique,” 4.

<sup>49</sup> Colin Hay, “State of the Art: Divided by a Common Language: Political Theory and the Concept of Power,” *Politics* 17, no. 1 (1997): 48.

<sup>50</sup> Frank Cunningham has made a similar point in reply to a different objection to talk of objective (Cunningham uses this term rather than “real”) interests. Responding to the concern that talk of objective interests could be used to justify totalitarian political projects, Cunningham writes that “nobody can deny that such ‘justification’ has been employed by totalitarians, but it does not follow that an account of objective interests cannot be given which avoids making them prescriptions for totalitarians.” Frank Cunningham, “Pluralism and Class Struggle,” *Science & Society* 39, no. 4 (December 1, 1975): 385–416.

<sup>51</sup> Meyerson, *False Consciousness*, 106.

<sup>52</sup> *Ibid.*

<sup>53</sup> *Ibid.*

<sup>54</sup> *Ibid.*, 109.

<sup>55</sup> Stephen K. White, “The Very Idea of a Critical Social Science,” in *The Cambridge Companion to Critical Theory*, ed. Fred Rush (Cambridge University Press, 2004).

<sup>56</sup> Katherine Cramer Walsh, “Putting Inequality in Its Place: Rural Consciousness and the Power of Perspective,” *American Political Science Review* 106, no. 03 (2012): 517.

<sup>57</sup> James H. Kuklinski et al., “Misinformation and the Currency of Democratic Citizenship,” *The Journal of Politics* 62, no. 3 (2008): 790–816; Suzanne Mettler, *The Submerged State: How Invisible Government Policies Undermine American Democracy* (University Of Chicago Press, 2011); Martin Gilens, “Political Ignorance and Collective Policy Preferences,” *The American Political Science Review* 95, no. 2 (June 1, 2001): 391.

<sup>58</sup> Gilens, “Political Ignorance and Collective Policy Preferences,” 383.

<sup>59</sup> *Ibid.*, 385.

<sup>60</sup> Stephen K. White, “The Very Idea of a Critical Social Science”; William E. Connolly, “On ‘Interests’ in Politics,” *Politics & Society* 2, no. 4 (September 1, 1972): 459–77.

<sup>61</sup> Steven Lukes, *Power: A Radical View*, 2nd ed. (Palgrave Macmillan, 2005), 38.