CORE
Metadata, citation and similar papers at core.ac.uk
Provided by PhilPapers

Dilemmas of Political Correctness

DAN MOLLER

University of Maryland

ABSTRACT

Debates about political correctness often proceed as if proponents see nothing to fear in erecting norms that inhibit expression on the one side, and opponents see nothing but misguided efforts to silence political enemies on the other.[1] Both views are mistaken. Political correctness, as I argue, is an important attempt to advance the legitimate interests of certain groups in the public sphere. However, this type of norm comes with costs that mustn't be neglected–sometimes in the form of conflict with other values we hold dear, but often by creating an internal schism that threatens us with collective irrationality. Political correctness thus sets up dilemmas I wish to set out (but not, alas, resolve). The cliché is that political correctness tramples on rights to free-speech, as if the potential loss were merely expressive; the real issue is that in filtering public discourse, political correctness may defeat our own substantive aims.

WHAT IS POLITICAL CORRECTNESS?

Political correctness, as I will understand it, is the attempt to establish norms of speech (or sometimes behavior) that are thought to (a) protect vulnerable, marginalized or historically victimized groups, and which (b) function by shaping public discourse, often by inhibiting speech or other forms of social signaling, and that (c) are supposed to avoid insult and outrage, a lowered sense of self-esteem, or otherwise offending the sensibilities of such groups or their allies. The concept, we should note, is one used by its enemies; dubbing something politically incorrect implies there is something worrisome or objectionable at work, though not necessarily that the po-

1. Earlier philosophical debates illustrate this. See, e.g., Friedman and Narveson 1995

litically correct option is wrong all things considered. But to avoid verbal disputes, let us simply take on board the language of "political correctness" and concentrate on the substantive merits of the doubts that are implicit in the pejorative tone.

According to this characterization, merely advocating for substantive policy changes is not itself a reflection of political correctness, except in a vague, by-association sense of the term. Criticizing someone for referring to an administrative assistant as a "secretary" is a manifestation of political correctness, but advocating for higher wages for assistants is not; insisting on trigger-warnings on syllabi or deleting offending material is again a form of political correctness, but arguing for rape-prevention security measures is not. Certain fringe environmental positions might themselves be loosely dubbed "PC" views, but I suspect this is only because the people who adopt such positions often advocate politically correct norms alongside. Symmetrically, it isn't politically incorrect to make a donation to fight gay marriage–though of course many would respond to doing so with outrage–but it is politically incorrect to write an op-ed making a careful, dispassionate argument against gay marriage. Political correctness thus isn't about private choices deviating from some norm; the notion doesn't refer to a distinctive personal morality, but to a system for moulding public discourse.

The norms involved are mainly, if not universally, negative and inhibitive, and many cases that initially seem positive are more complex, as when campus advocates urge multicultural curricula that move away from the Western canon, or argue for including more women or members of various other categories on syllabi, which doesn't initially seem inhibiting. But the underlying goal, even in these cases, is to avoid the sense that certain groups are marginalized or devalued because members of their group aren't represented in the canon or syllabus. What is being resisted in this kind of case is a certain *implication* that would otherwise inform public discourse, an implication that proponents of political correctness wish to eliminate. We might worry that this isn't what is happening when people simply point out that the Bronze Head from Ife, say, has intrinsic aesthetic merit that warrants study on par with comparable European art, but then this doesn't look much like an appeal to political correctness. The account, it's important to emphasize, isn't supposed to capture just any revisionary or vaguely "left" policy, but rather those with the flavor of responding to the sensibilities of marginalized groups by blocking an offending element. Making a case for the Ife head as envisioned above is an aesthetic argument motivated by independent, positive enthusiasm for the features of the work; appeal-

ing to the negative effects on the self-esteem of certain students when asked to study Phidias, Michelangelo and Picasso is an appeal to political correctness. The same goes for historians' arguments for the revision of inaccurately jingoistic textbooks versus those concerned with avoiding any implication that certain groups are inferior or that their grievances aren't worth addressing. (Of course, the distinction can be a difficult one to draw).

On this definition, moreover, it is significant that what makes a statement politically incorrect is not whether it in fact serves to promote the interests of certain people overall, but whether it threatens their public standing, as typically manifested in a sense of insult and outrage, or a lowered sense of self-esteem and inclusion. Notice, for instance, that no one is willing to retract judgments about what look like politically incorrect statements if they later turn out to promote the interests of the groups in question. If the president of the university says "Members of group G are underrepresented in field F because of unflattering trait T" this may well be judged politically incorrect, and that judgment won't change if it turns out that this was just what G needed to hear. The objective likelihood of advancing the cause of G is beside the point when it comes to political correctness. The air of political incorrectness is brought about by the insult itself, and thus the usual way of overcoming substantive criticism–by showing that the local harm of the insult was outweighed–is ineffective. Calling someone by some group-epithet does not become less politically incorrect if that turns out to be motivating and helpful to the individual, as insulting a friend at a tennis match ("Come on, you jerk!") can evade criticism if it turns out to be helpful overall. That is because the target of political correctness is the insult itself (along with the corresponding threat to the public standing of the group), not the overall interests of the people involved.

This might be resisted on grounds that there is evidence of "stereotype threat," i.e., that issuing politically incorrect statements like the university president's often negatively impact the actual performance of members of marginalized groups, simply by raising the salience of their group-membership and the perception that members are less good at a given task. (Subtly reminding test-takers that they are members of a marginalized group can cause them to perform worse than control groups that take the same test without a priming-cue.)[2] My characterization may then seem falsely to suggest that the concern is for something trivial-sounding ("not hurting people's feel-

2. For a summary and discussion, see Gendler 2011, 48-51. See also research on the possible impact of emphasizing native talent or brilliance, Leslie et al. 2015.

ings") whereas the ultimate concern is to prevent the real and documented damage that the relevant speech and behavior does. However, my claim is not that politically incorrect speech cannot have objectively damaging effects on others, or that such effects might not motivate politically correct norms. The idea is that what *makes* something politically incorrect is a certain kind of offense in virtue of undermining public standing, not that offending people in such a way need be trivial or that blocking such offenses might not have a deeper underlying motivation–just the opposite, as we will see.

As a final elaboration, my gloss emphasizes marginalized groups as the intended beneficiaries of politically correct norms. It is this aspect that leads me to differ with the economist Glenn Loury's otherwise searching analysis. Loury treats political correctness as a far more general phenomenon than I have, suggesting that as "an implicit social convention of restraint on public expression, operating within a given community" it applies to any such restraint, left or right, including, say, fascist censorship. (1994 p.430) The key for Loury is that political correctness, left or right, culminates in a kind of *self*-censorship through a feedback loop in which, first, there are sanctions for those who violate the communal norms, and then those who are willing nevertheless to risk such sanctions come to seem especially refractory.

Suspicious speech signals deviance because once the practice of punishing those who express certain ideas is well established, the only ones who risk ostracism by speaking recklessly are those who place so little value on sharing our community that they must be presumed not to share our dearest common values.(Loury 1994 p.436)

Loury emphasizes such examples as the German politician Phillipp Jenninger, who fell into disgrace after a speech that engaged rhetorically with the perspective of Nazi-era Germans, even though it was unambiguously clear that both the speech and Jenninger's prior life and work were devoid of Nazi-sympathies or anti-semitism. It is worth noting that after Loury's article was published, a Jewish leader gave the same speech in a synagogue in order to demonstrate what he rightly predicted would be the non-response.(*Die Welt* 1995) The worry, clearly, wasn't the substance of what Jenninger had to say but the *signal*, Loury would underscore, that is sent by a German politician (but not a Jew in a synagogue) being willing to take up, if only for rhetorical purposes, the perspective of Nazi-sympathizers *after* it has been established that taking up the Nazi-era point of view is taboo.

Loury is surely right about the impressive degree of self-censorship political correctness can achieve or demand. Standard examples examples include the white

Washington DC mayoral appointee who abjectly apologized to his colleagues who were outraged by his use of the word "niggardly" leading the mayor to accept his resignation. (*The Washington Post* 1999) But although self-censorship is a kind of ultimate victory for those wishing to eliminate some form of expression, actual censorship of various sorts is on the table as well. The same term has been the focus of college speech codes, as when a student objected to it in an academic setting even after its meaning and use by Chaucer were clarified ("It's not up to the rest of the class to decide whether my feelings are valid"). (*Reason* 1999) And of course it's natural to start with formal censorship in order to induce self-censorship. More importantly, I believe there is something distinctively *left* about political correctness, something connected to the concern for victims' groups. This may sound semantic–Loury and others could just announce their conception of political correctness is a bit broader than mine. But there are important differences between how and why various norms shaping public discourse originate and are enforced that are worth recognizing.

Take right-wing attempts to delegitimize opposition to war by suggesting dissenters are insulting "the brave men and women who fight on our behalf," or attempts to shape discourse concerning torture by insisting on Orwellian euphemisms like "enhanced interrogation." Such maneuvers are important to analyze in their own right, since they may work to inhibit speech in disastrous ways, often with outcomes far worse than anything to emerge from petty squabbles over how to refer to an office assistant. But that doesn't make the cases any less distinct. What motivates these right-leaning efforts, baleful though they are, is usually different from and nearly opposite to what motivates norms against questioning affirmative action or syllabi with too many dead white men. In a typical case, what motivates an effort to suppress dissent about war or torture is concern for security, not sympathy or feeling sorry for marginalized or oppressed groups. And the target of the norms is typically what is seen as a display of weakness rather than insults or offenses against the sensibilities of those marginalized, while the response tends toward contempt for the weak rather than outrage at the insult; accusations of disloyalty or spinelessness are more likely than those of being insensitive or cruel. There is a common danger that these attempts at molding discourse will backfire in ways we'll explore below, but these sub-differences, summarized in the table below, still result in an important overall difference in the character of what takes place. (Of course, these are just one set of possible differences; the right-list would differ for those motivated by a concern for individual liberty rather than security.)

	Left	*Right*
Motive	Sympathy	Security
Target	Offense	Weakness
Response to violations	Outrage	Contempt
Accusation	Cruelty	Disloyalty
Danger	Backfire etc.	Backfire etc.

Political correctness, then, is far from unique in trying to influence public discourse and in trying to compel people to speak or think in certain categories or terms. In discussing the problems associated with political correctness we are not singling out left-leaning concerns for special scrutiny. All kinds of social institutions, both left and right, shape which arguments get made, including libel and national security laws, and informal conventions that govern clubs or associations, each with their own profile of burdens and benefits. But political correctness *is* distinctive, and a distinctively left phenomenon, I want to insist. Those attempting to shape discourse on the right are rarely moved by feeling sorry for some group and rarely make corresponding objections focused on avoiding offense. And when they go wrong and undercut their own aims, as when their attempts to shape debate about a war turn out to undermine national security in the long run, they do so by exhibiting a characteristic series of mistakes that are distinct from those most common on the left.

LEGITIMATE ENDS

Political correctness is dismissed by its opponents as if it were either a bizarre and trivial insistence on redefining words, or else an insidious attempt to advance an ideology by silencing the competition. "Yes, but..." is not the typical response of those with reservations. Loury, for example, speaks of the "superficial moralism" of political correctness. This certainly applies to cases like "niggardly" that we can dismiss as childish. But it is less easy to dismiss the taboo on the N-word itself (herewith observed), harder still to dismiss certain taboos regarding racial science, and impossible, I think, to dismiss the underlying worries animating such strictures.

Consider as a historical example the response to the controversial book *The Bell Curve*, which claimed that there is such a thing as general intelligence, that IQ-differences are partly heritable, that they have a significant effect on social outcomes,

and that there are racial differences in average IQ as measured by a standardized test. That reactions to books like *The Bell Curve* really are manifestations of political correctness, not just anodyne scientific disagreement (though there was plenty of that as well), seems hard to deny. The clearest way of making such a diagnosis is to observe that the concern is overwhelmingly rooted in anxiety about offending or insulting a historically marginalized group. The key test is how similar research fares that either doesn't insult but rather tends to praise the group in question, or insults but not a group that we collectively feel sensitive toward. What is striking is that there is no widespread outrage about research into the cognitive *advantages* that Jews or certain Asian groups are sometimes said to enjoy (in scoring higher on average on certain kinds of tests), or the flipside to such research which is that various European groups are inferior in some respect. Here, again, there is plenty of scientific disagreement about the validity of the categorizations involved and the specific experiments or tests underlying the claims, but there isn't the collective shock of a taboo being breached and the accompanying outrage, what amount to public trials, excommunications, and so on.[3] Whatever the scientific flaws critics detected in *The Bell Curve,* political correctness is the only plausible explanation for the asymmetric treatment of parallel work that just isn't insulting or else insults groups no one feels much sympathy for.

So much critics of political correctness get right in a case like this. But they neglect the perfectly good *reasons* for cultivating and enforcing various politically correct norms. In this instance, the root concern is clearly that there exists a horrific record of violence and injustice directed toward African-Americans, as well as a record of promoting such violence by superficially respectable means (including racial pseudo-science), and enlightened moral thinking has thus converged on a *default norm against advancing ideas associated with the oppression or marginalization of African-Americans*. This is why leading responses to *The Bell Curve* focused on associating it with earlier instances of debunked racial science.(Gould 1996) Political correctness thus represents the evolution of public standards with the praiseworthy tendency to protect and promote the interests of historically oppressed groups. These standards work by introducing a *high barrier of entry* to those wishing to enter

3. In 1994 *The New York Times* published an entire series of articles denouncing *The Bell Curve*, a representative opinion piece being "The 'Bell Curve' Agenda" (Oct. 24, 1994). By contrast, the Times headline on research suggesting the heritability of high Jewish IQ scores was "Researchers say Intelligence and Disease May Be Linked in Ashkenazic Gene" (June 3, 2005). This piece, too, raised doubts about the thesis, but the difference in tenor is obvious from the headlines and the texts themselves. *New York Magazine* published an article entitled "Are Jews Smarter?", which also expressed substantive skepticism, but it is hard to imagine a symmetric article with an inverted title about some gentile group.

public discourse in a way that that threatens to undermine moral progress. By maintaining the norms, we acknowledge that such threats exist and that it is important to us collectively to signal to new entrants into public discourse that they must observe the norms carved out to protect the status of groups potentially under threat. And what is true in this case is true of many other examples of political correctness, such as censoring stereotyped depiction of Asians, the German anxiety over displays of sympathy for National Socialism, calls for including more women and other groups on syllabi, or suggestions that the poor are to blame for their plight.

In this respect, I am entirely in agreement with Richard Rorty that political correctness has made "the casual infliction of humiliation…much less socially acceptable than it was," and even that "encouraging students to be what mocking neoconservatives call 'politically correct' has made our country a far better place."(Rorty 1998 pp.81-2) There is no denying that norms to avoid insulting or otherwise attacking the status of women or gay people have brought huge benefits, and critics of political correctness who ignore them are simply mistaken. There are, to be sure, limits on the pursuit of these worthy ends, and inevitably disagreement about where to locate those borders. Barriers to entering the arena of public discourse can be higher or lower–at one end of the spectrum are minor conventions and taboos, the sense of collective shock when someone "dares" to utter certain things. At the other end are explicit laws, say prohibiting hate-speech, which may themselves be narrowly or very broadly defined. Some Canadian jurisdictions, for example, have made it a human rights violation to make any "vexatious comment" known to be "unwelcome by the individual or class" on grounds that include "political belief," which one might reasonably fear as absurdly overbroad.(Northwest Territories Human Rights Act, sec. 14(2), as of 2015.) One may acknowledge the legitimate ends of political correctness without endorsing any and all barriers to public discourse. Correspondingly, we may resist sticks-and-stones maxims suggesting anything-goes in public discourse while symmetrically resisting attempts to shape public discourse by certain agents (such as the state) or to certain degrees (refusing to hire anyone who says anything vaguely "un-PC"). These may, after all, produce costs or pose dangers in their own right.[4] And of course any particular instance of political correctness may be wrongheaded or petty, just as individual applications of patriotic norms. We must not, as Rorty ultimately does, lose sight of the potential drawbacks to political correctness so as to

4. See Waldron 2012 for an argument that the state should in fact pursue heavy-handed tactics like hate-speech laws in pursuit of the sort of legitimate ends I have been acknowledging.

arrive at a reasonable estimate of what, all-in, we gain and suffer, in upholding these norms.

These, then, are legitimate ends for political correctness. Political correctness in itself needn't be mistaken in its fundamental aspirations. Proponents of PC-norms aren't confused to think that racial pseudo-science has had enormous, damaging effects in the past; they aren't mistaken to regard any revival of racial science as potentially disastrous and in any case accompanied by huge costs. It is not true that opposition to any such revival is (or need be) rooted in mere "superficial moralism," and there are good reasons for maintaining collective default-norms that signal certain kinds of discussion out of bounds in the normal course of things.

There are two wrinkles in this story that bear mentioning, however. One is that politically correct norms have a distinctive *content* that makes emphasis on language inevitable. The whole point of such norms, as I have described them, is to generate a set of default-presumptions that those participating in public *discourse* are expected to observe in order to ward off threats to a certain kind of moral progress, and so naturally terminology and word-choice features prominently in the marshaling of such norms. This can then give rise to the absurd cases already noted that often revolve around what really are morally superficial *applications* of reasonable norms. Norms pertaining to language-use are perhaps especially liable to misuse in ways that will strike some as preposterous since they inevitably implicate what can always be ridiculed as "mere" labels. Relatedly, we noted that political correctness concerns offense and sensibilities, not the objective interests of those involved. It might seem surprising that the norm to evolve was one that focused on offense and not simply on promoting whatever was in the people's objective interest. But this is again similar to other norms, like love of country. In both cases there is a core goal of promoting the interests of some entity, but part of this is taken to involve discouraging insults and other threats to the publicly recognized status of the people or thing in question. Failing to acknowledge the values in question by a lack of reverence or deviance from certain standards are thus punished, even when what is at stake seems superficially to be only symbolic. Political correctness is one face of a deeper concern for the oppressed comparable to the dimension of patriotism associated with denouncing insults to country.

DILEMMAS

There is nothing wrong with promoting a presumption that historically oppressed or marginalized groups should not be insulted or subjected to discourse threatening to undermine their status, and it is puzzling that critics of political correctness seem frequently unwilling to acknowledge its legitimate ends. That leaves the door open to a second kind of criticism, the misguided application of the relevant norms, but whatever the damage to individuals losing their jobs or being publicly anathematized, it cannot be said that mere misapplication of values raises interesting philosophical problems. It is rather a third kind of problem with political correctness that should anchor our attention, the problem of *conflicts* among values, whether between those associated with political correctness and other things we care about, or even internal conflicts within the former. We can enumerate several different kinds of dilemma-engendering conflicts.

Orwellian discourse: One kind of conflict occurs when politically correct norms lead to the kind of abuse of language that Orwell criticizes in "Politics of the English Language." We noted earlier that the petty misapplication of language-norms isn't worth making a fuss over, but as Orwell points out, the vague and imprecise use of terms like "fascism" can come to serve as a "defense of the indefensible."(1946 p.162) Contemporary versions of this on the right are easily recognized, as when a kill list becomes a "disposition matrix," but political correctness seems to involve a similar tendency. Loury draws the connection to phrases like "disadvantaged minorities" (nowadays the term would be "diversity") which he says is "used in educational philanthropy circles when the speaker really means 'non-Whites, excluding Asians.'... Such linguistic imprecision impairs analysis. But that is often its purpose," among other reasons, he suggests, because announcing that a scholarship was to be offered to "non-whites excluding Asians" would, by its very accuracy, render the proposal impossible.(Loury 1996 p.447) Another policy-shaping example is the increasing tendency to reject official government terms like "illegal alien" in favor of "undocumented immigrant" or even "undocumented citizen," with the implication that refusing to do so implies reactionary or hateful views. These campaigns aren't just the one-off ideas of random individuals; the phrase "undocumented citizen" is encouraged by administrators at a major state university in the United States, and others urge that

the statement "America is a melting pot" constitutes a form of "microaggression."[5] Regardless of what the right immigration policy is, and notwithstanding the legitimate interest in avoiding various forms of marginalization, this kind of discourse once again "impairs analysis." "Undocumented immigrant" is meant to make it harder to focus on the fact that there are laws and procedures governing entry to the country that were flouted by the persons in question, while the Orwellian "undocumented citizen" seeks to present a political aspiration as a *fait accompli*. To the extent that we recognize both the legitimate ends of political correctness and the undesirable effects that Orwell drew our attention to, we should see these as cases that present a dilemma.

Causal structures: More substantively, fears of politically incorrect stereotyping threaten to subvert our understanding of the world even without Orwellian word-games, as when there is resistance in the public sphere to the suggestion that a stereotypical trait is causally implicated in some negative outcome. The stereotype of deference to authority in many East Asian (and other) societies and its role in causing accidents is an example.[6] When Korean airline flight Asiana 214 crashed in San Francisco, the suggestion was made that such deference made a difference, as the pilot was relatively junior and was being supervised by an instructor, possibly leading the pilot to be reluctant to assert the need for a go-around. This suggestion was in turn widely derided for succumbing to cultural stereotypes. *Atlantic Monthly* author James Fallows, for example, introduced the claim under the heading "Confucius in the cockpit," alongside a comical depiction of the sage, and noted that he was, "highly skeptical of this whole line of thinking...If an (apparently) mishandled approach shows something about Korea–or East Asia, or Confucius, or rote-learning systems–then what do we make of the many thousands of Asian-piloted flights that land smoothly and safely throughout Asia every single day?" (Fallows 2013) Needless to say, the safe landings Fallows refers to aren't reasons to discount or mock the suggestion that a trait exhibited with a higher prevalence in some cultures

5. "Undocumented citizen" occurs for example in an official publicity campaign of the University of Maryland, which places "illegal alien" alongside expressions like "retarded" and "no homo." http://thestamp.umd.edu/multicultural_involvement_community_advocacy/programs/inclusive_language/phrases, accessed 2/3/15. The "microaggression" point is from "Tool: Recognizing Microaggressions and the Messages they Send," part of the materials for a leadership seminar sponsored and extensively promoted by UCLA. http://www.ucop.edu/academic-personnel-programs/programs-and-initiatives/faculty-diversity-initiatives/faculty-leadership-seminars.html, accessed 7/20/2015. For more details, see Volokh 2015.

6. For a review of how accurate stereotypes are, see Lee et al. 1995. The stereotype that stereotypes are generally wrong is itself dubious, as the authors point out. For an accessible historical survey focusing on Asian aviation safety, see Gladwell 2008, ch. 7; I focus on a more recent example.

than others might contribute to the explanation of an accident in this case. What appears to be at work in this writing as well as in other public denunciations of the hypothesis is anxiety about reinforcing stereotypes. The important thing for our purposes isn't whether authority-deference actually *did* play a role, only that politically correct norms threaten rational analysis of the cause of a plane crash, assuming that public ridicule counts as a cost those analyzing such crashes must reckon with. As it happens, "Interviews with pilots indicate that Korean culture may have played a role in the crash...Captain Lee told investigators that any of the three pilots on the plane could have decided to break off the approach, but he said it was 'very hard' for him to do so because he was a 'low-level' person being supervised by an instructor pilot."(*The New York Times* 2013) The NTSB report states that "the PF's [pilot flying's] deference to authority likely played some role in the fact that he did not initiate a go-around."(NTSB/AAR-14/01, 92)

Against this, the fact that there are sources to cite discussing the role of cultural differences in accidents may seem to undercut the idea that there is some politically correct taboo surrounding the subject. But politically correct norms are graded–some topics are widely off-limits in public discourse, but others merely get subjected to "heightened scrutiny." These introduce a filtering effect. The thought isn't that it is impossible to discuss publicly the arguments involved, but knowing one will be subjected to moralized criticism introduces an initial barrier serving as a partial filter on public speech. Similarly, no one thinks it was impossible to criticize the Iraq war, but patriots seeding suspicion of dissent in effect raised the barrier to entrants to the public discussion. And it's worth observing in passing that the barrier introduced by stereotype-aversion extends beyond political correctness strictly speaking, to avoiding stereotypes more broadly, again with worrying effects. Neanderthal research, at least any that informs public discourse is inevitably along the lines of, "surprising new study upends stereotype that Neanderthals were dim-witted." This in itself might seem to reflect a random piece of scientific progress. The trouble is that it is difficult to imagine an article title, let alone a newspaper headline, along the lines of "Neanderthals: as dumb as we thought." Such research would thus need to overcome both bias in favor of novelty and the quasi-politically correct bias against saying anything nasty about the underdog (whom the extinct presumably exemplify).

Backfire: Other conflicts are internal to the concerns the community has for those marginalized, particularly conflicts arising between the public-facing desire not to insult or offend, on the one hand, and the substantive concern actually to advance

people's interests on the other. We see this in the case of patriotism when jingoistic zeal interferes with frank and open attempts to improve the life of the country. Dissent in war is the obvious case, but there are many others, as when critics on the right refuse to accept "revisionist" histories that attempt to wrestle with an ugly past so as to improve national culture, or when national pride leads to a denial that core values are being undermined by various policies. These conflicts represent a set of norms backfiring against those who apply them so that the core-values the norms emerged from are actually undermined as a net result. Politically correct backfiring includes pressure for trigger warnings in courses and attendant pressure to leave off "sensitive" materials from the syllabus that may upset students who have been traumatized. This has the predictable consequence that instructors are less likely to teach material relevant to topics like sexual violence, with the unintended consequence of decreasing knowledge of relevant law that might be used to protect women: "asking students to challenge each other in discussions of rape law has become so difficult that teachers are starting to give up on the subject" leading instructors to omit "rape law in their courses, arguing that it's not worth the risk of complaints of discomfort by students."(Suk 2014)

The so-called mismatch hypothesis concerning American-style affirmative action furnishes another example of the costs of political correctness. (For an overview, see Sander and Taylor 2012.) Once again, my point does not turn on whether the empirical claim is true; as with any social science work, there is bound to be some controversy and my goal isn't to establish the validity of a particular scientific claim. But according to a significant body of research, affirmative action does immense damage to the "beneficiaries" of the program by tending to shift students from academic environments in which they might well flourish toward harsher, more demanding environments for which they may not be as well prepared, and in which they consequently do worse. The main problem appears to occur not at the higher echelon of elite institutions, but in somewhat less selective schools, who as a result of an under-supply of suitable students are left with fewer and fewer candidates, as those there are get scooped up by the more elite schools. Diversity-pressure at higher echelons, in other words, is said to have disastrous consequences at lower echelons that fill their ranks with candidates who would benefit more at less selective institutions. These benefits are said to include better grades, greater learning, better bar-exam results, greater likelihood of going into the sciences, and better careers. For example, following the natural experiment introduced by California's proposition

209 barring affirmative action, the number of African-American and Hispanic freshmen who went on to graduate in four years rose 55% and the number who went on to get a STEM degree in four years rose 51%. (Sander and Taylor 2012 p.154)

Obviously, it is reasonable to wonder whether the benefits of being admitted to an elite institution outweigh the benefits of ending up at a less elite but better "fit" institution, or how common such dilemmas really are, and so on. But a body of academic research had accumulated that this might be so by the early 2010s. According to this work, there were large net benefits to attending a school for which students were well prepared academically instead of a fancier school in which they were more likely to struggle. This research had a reasonable hypothesis as its target, was performed by multiple, well-respected faculty at prominent institutions and was published in serious, peer-reviewed journals. Nevertheless, there was (and is) a profound resistance to taking any of this research seriously, despite the fact that it purported to show that a set of policies was backfiring so as to cause the community to fail to achieve its own aims.[7] At Duke, research showed that non-Asian minorities tended to self-select out of the hard sciences because of poor performance as a consequence of mismatch, but instead of prompting corrective action, school officials reacted with lukewarm affirmations of academic freedom, and the comment that "We understand how the conclusions of the research paper can be interpreted in ways that reinforce negative stereotypes." (Lange et al. 2012) Any sense that there was powerful evidence that our policies might be irrational (in the formal sense that they caused us to act contrary to our own self-given aims) was (and still is) almost entirely absent. And at the second order, the authors received the scathing denunciations characteristic of political correctness, including at campus protests, signaling that research into what would promote the substantive interests of historically oppressed groups would not be tolerated if the results conflicted with norms against insult and offense.

I am not claiming that universities and public intellectuals were wrong, all things considered, to ignore or deny this research. We have seen that there are legitimate reasons to adopt a default norm of straight-arming ideas tending to insult the status of marginalized groups. The point is rather that doing so comes with *costs*, setting up a *dilemma*. My central contention isn't that we ought to do away with the supposed superficial moralism of political correctness, but rather that we ought to focus instead

7. For a striking illustration, see the high-profile public debate on the subject "Affirmative Action on Campus Does more Harm than Good" (Intelligence Squared US), widely available online.

on the dilemmas political correctness introduces, and face up to the costs incurred in being gored on either horn.

NORM-DEPENDENT RESPONSES AND REVERSE-HYPOCRISY

It is difficult to tally up or even to compare the costs of having or not having politically correct norms, but it is clear that both can be high. To dwell on the costs of enforcing them, in a high-cost scenario they can lead to what Timur Kuran calls widespread "preference falsification" in which what people believe in private becomes increasingly detached from what is spoken in public, which in the case of East-bloc communism made it impossible to discuss pervasive dysfunctions urgently requiring reform. Worse, Kuran identifies an "intergenerational process through which the unthinkable becomes the unthought," making such dysfunctions unidentifiable even if they could be discussed. (Kuran 1995 ch.13 and p.186) Alternatively, Loury points out that preference falsification can lead to polarization, by a process analogous to Gresham's law, whereby the bad money (extreme opinion consonant with politically correct ideology or else violently opposed to it) drives out the good (moderately heterodox opinion), and so comes to dominate what circulates in public. (Loury 1996 pp.435-6)

This is the high-cost scenario for political correctness. We can illustrate it in the arena of distributive justice, which offers ample scope for the relevant norms. Attributing poor social outcomes to factors *external* to the person (to society, the state, etc.) sounds "nice," since we don't feel like we're blaming the underdog for their already unpleasant position; attributing them to factors *internal* to the person (e.g., to poor choices) sounds "mean" and is likely to trigger charges of "blaming the victim." This makes the latter less politically correct. And on the international stage, claiming that poor countries are in part worse off due to endogenous factors like institutions or culture similarly has a un-PC quality to it that blaming multinational corporations or the rich countries does not. Since there is a long and ugly history of rich countries invading poor countries and an even longer history of richer citizens taking advantage of poorer citizens in politics, law and business, it isn't unreasonable to accept a norm discouraging theories threatening to undermine the status of the poor. But against this, if it turns out that people are capable of significantly influencing social outcomes in the course of educational, fertility or work decisions, and that absent these, statist policies are likely to be ineffective, it will be disastrous for such facts

not to inform public discourse, or for them to face ridicule. (Similarly in the international variant.) The high-cost scenario in these cases, then, is one in which there is a widespread belief that, say, social pathologies play an important role in explaining bad social outcomes, but there is reluctance to discuss that belief; or in which the thought doesn't seem even a live option to many (it's "unthought"); or that damaging polarization sets in because those with moderate views face penalties for voicing their opinions in the public arena ("Most people can exercise significant influence on whether they end up poor and should be criticized for not doing so positively, but of course violence-prone slums or abandoned rural areas are another story").

This high-cost scenario seems to me more than overblown fear-mongering, though it would be difficult to establish the extent to which it is or may be realized. Instead, let me make two specific suggestions about how to think about the costs on either side of the ledger. On the one side, there is a curious problem that arises when one is concerned to promulgate norms so as to avoid insult or offense, but those very norms play a role in shaping the nature of the insult or offense. In Germany Americans are sometimes referred to as "Amis," short for Amerikaner (not the French term). Suppose a headmaster notices that this expression is used of a small minority of expatriot students who are sometimes bullied. In sympathy, the headmaster forbids what he sees as a possibly condescending use of "Ami," insisting instead on the full "Amerikaner," and goes on to lecture students on how to treat their foreign guests. Is this the proper *remedy* for oppressed ex-patriots? It may be, and the old-fashioned approach of telling the ex-pats to buck up and pay no mind to the rotten kids may turn out to be ineffective or cruel. But the danger is that the headmaster's sympathetic norm itself sensitizes the ex-pats to what they formerly paid little mind to, but now interpret as major offenses that they ought to dwell on, talk about, feel traumatized by, and so on. A parent might reasonably judge the headmaster's approach a mistake, once the subtle point about the feedback loop implicit in such norms is recognized. In promulgating norms designed to benefit marginalized groups we both help and hurt them.

There is ample empirical evidence for this idea in relation to serious trauma. Victims of sexual abuse and combat veterans fare worse the more they see their traumatic event as a central, defining moment; norms tending to *downplay* the importance of the event would thus be expected to help. (Robinaugh and McNally 2011) A *New England Journal of Medicine* piece on victimhood and resilience points out that immediate counseling after trauma, which tends to highlight that the victims *are* victims

who should be *expect* to feel traumatized, often seems to make things worse, increasing the likelihood of mental disorders.[8] Given all this, we must be cautious in thinking about how to assess the costs of having or not having some norm that superficially promotes some victimized groups' interests; assessing the overall effects is far from straightforward.

On the other side of the ledger, an important metric to pay attention to (and which social scientists could attempt to measure) is the prevalence of reverse hypocrisy. This is the practice of applying high standards in one's private life, especially toward one's children or other loved ones, while publicly promulgating low standards for others, either explicitly or by withholding public criticism. Reverse hypocrisy is evidence of something like Kuran's preference falsification. The savvy communist party member publicly signals agreement with low standards for productivity, urging that the state should provide for everyone's needs without anyone doing "extra" work not officially assigned to him; but privately he urges his kids to work long hours in the informal sector and to accumulate savings. Examples closer to home include private insistence on personal responsibility in domains like fertility decisions or work ethos, while ignoring or even mocking these as public norms. In this sense, it's reverse hypocrisy to make it clear that you expect your children to make sensible decisions about family and to work hard in school whatever the excuses they are tempted to make, while criticizing or lampooning old-fashioned sounding public norms to the same effect.

It might be objected that there are substantive reasons to uphold different standards in the public arena than in the personal. A liberal tolerance for differences or even just politeness might seem to dictate as much, and of course if one judges that others are less fortunate in their capacity for making the relevant discriminations, or are less well positioned to act on them, it may seem inappropriate to uphold what would be unreasonable standards. Arguments from liberal toleration or etiquette are less persuasive, though, when the stakes include the wellbeing of someone else's family. And the view that *we* or those close to us can adhere to high standards that will promote our wellbeing but *they* cannot, has a worrying ring of condescension. Short of extreme circumstances, many successful families simply will not tolerate children doing poorly in school (let alone not finishing), making poor fertility decisions, or failing to work. But many of the same people are reluctant to assert these as public norms or to issue criticisms based on them. Since asserting such norms would

8. Wessely 2005. For philosophical reflection on our propensity to underestimate resilience in the face of trauma, see Moller 2007.

often involve criticizing marginalized groups–those on the receiving end of such criticism would almost by definition be worse off–this sort of reluctance looks like a good measure of political correctness, and its prevalence would be a useful barometer of what I call the high-cost scenario.

IS POLITICAL CORRECTNESS A MYTH?

Given the many examples cited up to this point, it may seem surprising that some have doubted whether political correctness exists at all, at least to any significant extent. But writers have in fact expressed two kinds of doubt along these lines.[9] One is the denial that there are socially significant instances of public speech being shaped in the ways I have outlined, so that it is denied that public speech about race or gender say is subject to anti-marginalizing norms to any significant extent. The other form that denial takes is insisting that, while there really are norms informing this discourse, these aren't motivated by the kinds of concerns I have singled out. If it turns out that "political correctness" is just the neutral struggle for truth and justice that the recalcitrant wish to rebrand and demonize, then perhaps the phenomenon is, once again, a kind of myth. (There is a tendency to combine these two forms of denial, but notice that they are inconsistent and so we should stick with one or the other.)

Presumably we can demonstrate that X exerts a non-trivial influence on public discourse by showing that the discussion of major social institutions is in part shaped by X. It should be sufficient to dispel doubts about the existence of political correctness that we demonstrate non-trivial instances of public discourse being subjected to its influence. And it seems to me that we have seen ample evidence that this condition is satisfied. Immigration laws are important social institutions. So are university admission policies and government investigative agencies. In each instance, as we have seen, there have been non-trivial exertions of influence by powerful entities such as university administrators and members of the press in order to shape discussion of the relevant issues. That just is political correctness, provided it is motivated in the way I have defined the phrase. There may be reasonable disagreement about how much political correctness there is (compare: "Exactly how much jingoism or

9. They include an anonymous reviewer, Feldstein 1997, Wilson 1995, and Fish 1994, in varying degrees. The latter three are responding to culture-war polemics from the right (e.g., Feldstein 1997, 116-120, Wilson 1995, 10-15, Fish 1994, 53-79), not careful analysis like Loury's, and so their doubts should perhaps be taken in that light.

sexism is there?"), or whether it is a major concern in the grand scheme of things ("Compared to all the other problems in the world, how bad is jingoism or sexism?"). But we shouldn't move from views about how important, exactly, political correctness is to denying its existence outright.

Alternatively, there is the suggestion that what is deemed political correctness is just a pejoratively described, politically neutral attempt to fight for truth and justice. But on reflection, this too succumbs to the evidence assembled up this point. The problem is that the entities seeking to influence public discourse seem specifically motivated to defend historically oppressed or marginalized groups and not other kinds of groups. Thus, if the objections to works like the *The Bell Curve* were rooted in neutral concerns about shoddy science, we would expect to see symmetrical concern for claims about inferiority and superiority, and among those that are supposedly inferior, similar concern for historically oppressed and historically dominant groups. But as noted earlier, that is not the case. Findings of high IQ scores among Ashkenazi Jews do not produce the same degree of social anguish and institutional ostracism, and no one is worried by the implication that gentiles are inferior. This doesn't show the scientific objections aren't correct–I am not drawing the invalid inference that because outrage was triggered by political correctness, therefore charges of shoddy science are wrong or can be dismissed. Let us just assume all of the scientific objections were correct. The point is that there's a distinctive concern for the status of oppressed and marginalized groups at work here, not that such motivations cannot serve to uncover the truth. A similar symmetry-test applied to the other cases discussed produces similar results. University officials urging us to refer to illegal aliens as "undocumented citizens" make similar suggestions concerning other marginalized groups, we noted, not concerning dominant groups who might be deemed mislabeled. Those concerned about the role of stereotypes in causal explanations aren't symmetrically concerned to stamp out stereotypes about dominant groups. Nor is any of this surprising. It would, if anything, be strange if a well-meaning public failed to have some norms about public discourse concerning historically marginalized members. As long as the public does, we should expect these kinds of asymmetric norms which, as I have argued, have legitimate ends but also pose difficult dilemmas.

Pressing on such asymmetries may seem misguided if there are real differences between the cases. Discourse that suggests that historically marginalized people somehow deserve to be marginalized is obviously harmful in a way that insulting dominant groups is not. It is no wonder that we respond differently to these differ-

ent cases, we may think, which involve harms that can hardly be compared. This, however, is to make my point for me. The differences involved are precisely those that make for political correctness. To say that "it's different" when what is at stake is the public standing of a group that has been persistently wronged in the past is just to say that it's different when it's politically incorrect. This is what I have tried to argue all along.

REFERENCES

Fallows, James. 2013. Confucius in the Cockpit, and Other Items to Read, and Ignore, on Asiana 214. The Atlantic. [Available at: http://www.theatlantic.com/technology/archive/2013/07/confucius-in-the-cockpit-and-other-items-to-read-and-ignore-on-asiana-214/277703/, accessed 1/14/15].

Feldstein, Richard. 1996. *Political Correctness: A Response from the Cultural Left* (Minneapolis: University of Minnesota Press).

Fish, Stanley. 1994. *There's No Such Thing as Free Speech* (New York: Oxford University Press).

Friedman, Marilyn and Jan Narveson. 1995. *Political Correctness: For and Against* (London, Rowman and Littlefield).

Gendler, Tamar. 2011. "On the Epistemic Costs of Implicit Bias," *Philosophical Studies* 156, 33-63.

Gladwell, Maclcolm. 2008. *Outliers* (New York: Little Brown).

Gould, Stephen Jay. 1996. *The Mismeasure of Man* (New York: WW Norton).

Kuran, Timur. 1995. *Private Truths, Public Lies* (Cambridge: Harvard University Press).

Lange, Peter et al. 2012. "A message from Administrators Regarding New Study." *The Duke Chronicle*, January 18. [Available at: http://www.dukechronicle.com/article/2012/01/message-administrators-regarding-new-study, accessed 28 June 2016].

Lee, Yueh-Ting, et al. 1995. *Stereotype Accuracy: Toward Appreciating Group Differences* (Washington DC: American Psychological Association).

Leslie, Sarah Jane et al. "Expectations of Brilliance Underlie Gender Distributions Across Academic Disciplines" *Science* 347 no. 6219 (2015), 262-265.

Loury, Glenn. 1994. "Self-Censorship in Public Discourse," *Rationality and Society* 6, 428-461.

National Academies. 2008. *Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence*

The New York Times. 2013. "Pilots in Crash Were Confused About Control Systems, Experts Say," December 13. [Available at: http://www.nytimes.com/2013/12/12/us/asiana-airlines-crash-san-francisco-airport.html, accessed 1/22/15].

Moller, Dan. 2007. "Love and Death," *The Journal of Philosophy* 104, 301-316.

Orwell, George. 1946. "Politics and the English Language." In his *Collection of Essays* (New York: Harcourt Brace).

Reason. 1999. "Cracking the Speech Code." July. [Available at: http://reason.com/archives/1999/07/01/cracking-the-speech-code, accessed online 1/14/15].

Robinaugh, Donald and Richard McNally. 2011. "Trauma Centrality and PTSD Symptom Severity in Adult Survivors of Childhood Sexual Abuse." *Journal of Traumatic Stress* 24, 483–486.

Rorty, Richard. 1998. *Achieving our Country* (Cambridge: Harvard University Press).

Sander, Richard and Stuart Taylor. 2012. *Mismatch* (New York: Basic Books).

Suk, Jeannie. 2014. "The Trouble with Teaching Rape Law." *The New Yorker* Dec. 15. [Available at: http://www.newyorker.com/news/news-desk/trouble-teaching-rape-law, accessed 28 June 2016].

Vallone, Robert, Lee Ross, and Mark Lepper. 1985. "The Hostile Media Phenomenon." *Journal of Personality and Social Psychology* 49, 577-585.

Volokh, Eugene. 2015. "UC Teaching Faculty Members not to Criticize Race-based Affirmative Action, Call America 'Melting Pot,' and More." *Washington Post* 6/16. [Available at: https://www.washington-post.com/news/volokh-conspiracy/wp/2015/06/16/uc-teaching-faculty-members-not-to-criticize-race-based-affirmative-action-call-america-melting-pot-and-more/, accessed online 28 June 2016].

Die Welt. 1995. "Keiner hat etwas gemerkt." January 12. [Available at: http://www.welt.de/print-welt/article664397/Keiner-hat-etwas-gemerkt.html, accessed online 1/14/15].

The Washington Post. 1999. "Williams Aide Resigns in Language Dispute." January 27. [Available at: http://www.washingtonpost.com/wp-srv/local/daily/jan99/district27.htm, accessed online 1/14/15].

Waldron, Jeremy. 2012. T*he Harm in Hate Speech* (Cambridge: Harvard University Press)

Wessely, Simon. 2005. "Victimhood and Resilience." *New England Journal of Medicine* 353;6, 548-550.

Wilson, John. 1995. *The Myth of Political Correctness* (Durham: Duke University Press).