Racial Figleaves and the Shifting Boundaries of the Permissible

Jennifer Saul

We live in a time when 'racist' is one of the worst insults around. Very few people openly profess a view that they label 'racism', and those that do are very much marginalized. (Even denizens of the alt-right will often call their views 'racialist' rather than 'racist'.) And yet, as we all well know, we are devastatingly far from eradicating racism. But even those well aware of racism's power have been surprised by the overt racism of the Republican Presidential Primary in 2016. Despite, or perhaps because of this, Trump has triumphed to become the Republican presidential candidate. What does this mean? Is it becoming acceptable to be *openly* racist? How much worse, one finds oneself thinking, will things become?

This paper is an examination of how this political moment could come to be: how a shift can take place in which once unutterable sentiments become increasingly utterable. In particular, it's an examination of a particular device that, I will suggest, facilitates it—one that easily goes unremarked, but that in fact very much deserves scrutiny. I call this a 'figleaf', and the variety that is my focus here a 'racial figleaf'. A racial figleaf is an utterance[1] made in addition to an otherwise overtly racist one, that serves the function of calling into question the racism of the speaker and the utterance. I use the term 'figleaf' because it is an utterance that provides a small bit of cover for something that is unacceptable to display in public.[2] In what follows, I first outline the state of affairs that makes racial figleaves an appealing and useful technique to deploy; then I turn to a discussion of how these figleaves function in the changing of our conversational standards.

[1] Sometimes a figleaf might also be something other than an utterance, like a symbol or even a person: arguably, non-white UKIP spokespeople serve as human figleaves.

[2] One consequence of this is that which things figleaves will be used for will vary from culture to culture and time to time, depending on which things are considered unacceptable to show in public.

But before I go any further, a terminological note is in order. I am casting this discussion in terms of 'racism'. I mean this in a broad sense, encompassing prejudices of nationality or of religion as well as prejudices that map more clearly onto traditional conceptions of race. This is a bit of an oversimplification, as I think there are important differences between these sorts of prejudice.[3] However, I don't think anything turns on it for the purposes of the present paper.

1. Background: Norms, Resentments, and Processes of Change

1.1 Mendelberg, Implicit and Explicit Appeals

Although there have been some recent shifts (more on this soon), it still seems correct to say that the overwhelming majority of white Americans do not currently self-identify as racists, and they would be horrified by the thought that they are guilty of racism. As Tali Mendelberg (2001) notes, it was not always thus: Prior to the Civil Rights Movement, one could run and win at a national level on a platform that included claims of white supremacy and policies of legal segregation of the races. However, Mendelberg, argues, things changed. At the time of her writing (2001, 2008 a, b), American political discourse was governed by what she calls the Norm of Racial Equality. And yet, she noted, the majority of white Americans showed remarkably high levels of what psychologists call "racial resentment", measured by level of agreement with claims like "Irish, Italian, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors." (Tesler and Sears 2010: 19)

Mendelberg argued that this situation gave rise to a very specific sort of political messaging. She drew a contrast between *explicit racial appeals*—defined as those which use explicit racial vocabulary; and *implicit racial appeals*—ones which were far subtler, alluding to race either via images (most famously, the Willie Horton ad discussed in her 2001) or code words like "inner city" or

[3] I discuss these issues in more detail in my (in progress) paper "'Immigration' in the Brexit Campaign: Dogwhistle terms in complex contexts".

"welfare" (Horwitz and Peffley 2005). Implicit appeals—what I elsewhere call 'covert dogwhistles'—very effectively activate racial attitudes without a voter's awareness, while explicit appeals trigger self-monitoring and are therefore less effective. This is why the Willie Horton advertisement— an implicit appeal— caused racially resentful voters to support George HW Bush. And it is also why, Mendelberg argues, Jesse Jackson's criticism of this ad as racist caused this effect to dissipate (even though his criticism was treated as utterly misguided by mainstream media).[4] At this point, the appeal ceased to be implicit and became explicit, which rendered it ineffective. Through the 1990s and early 2000s substantial evidence accumulated for Mendelberg's contrast between the workings of implicit and explicit appeals.

Recently, the stark contrast between implicit and explicit appeals seems to have dissipated. There were initial indications of this in work by Huber and Lapinski (2006, 2008). But Mendelberg responded, pointing to potential design flaws in their experiments. More recently, however, Valentino et. al. (2016) appear to have demonstrated, in experimental work conducted 2010-2012, that explicit and implicit racial appeals can be equally successful. Their work is not susceptible to Mendelberg's criticism, and if they are right it is now clear that an explicit racial appeal can work just as successfully on racially resentful subjects as an implicit racial appeal. They conclude that "the power of racial attitudes in mainstream American politics is no longer dependent upon the ways in which race is discussed" (2016: 6). This change is largely attributed to the election of Barack Obama, and the subsequent widespread belief that racism is no longer an issue.

1.2 The Norm of Racial Egalitarianism

One might think this means that the Norm of Racial Egalitarianism is no longer. Certainly, the rise of Donald Trump has led to much speculation along these lines. But I think that this is premature, and that there is evidence of this in the

[4] I discuss these further in my forthcoming "Dogwhistles, Political Manipulation, and Philosophy of Language".

details of Trump's utterances. First, I think it is worth noting that Valentino et. al. have not demonstrated a conclusion quite as strong as that which they assert in the quote above. They have demonstrated that making race explicit no longer nullifies the impact of racial resentment on candidate and policy preferences, a deeply important finding. However, it's a further step from this to the idea that anything goes—that "the power of racial attitudes in mainstream American politics is no longer dependent upon the ways in which race is discussed" (2016: 6). And I don't think we should make this step. In fact, it seems quite likely to me that the Norm of Racial Egalitarianism is still in force.

To see this, let's think a little more about the form that the Norm of Racial Equality needs to take. The majority of white Americans were, after all, said to accept the Norm of Racial Equality while displaying high levels of Racial Resentment. Any remotely demanding Norm of Racial Equality would prevent one from endorsing items like "Irish, Italian, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors." (Tesler and Sears 2010: 19). The fact that the Norm is compatible with the endorsement of statements like these means that the norm must be a rather thin one. Mendelberg never gives us a statement of the norm, but she does state that it involves "the notion that racial inequality was an immoral principle"; "opposition to [white supremacy] and to the legal segregation it defended"; and discrediting of "the idea of biological inferiority". She notes that under the Norm, "neither citizens nor politicians want to be perceived or to perceive themselves as racist" (18). My working hypothesis is that the Norm can be understood as taking the very simple form **"Don't be racist"**. Adherents then apply their own understandings of what is required to not be racist—and in many cases, this is not very much.

Here, it is useful to look at the work of Jane Hill, who describes that she calls the "folk theory" of white racism. According to Hill, a key component of this is the Ideology of Personalism, which holds that: "racism is entirely a matter of individual beliefs, intentions, and actions" (2008: 6). If we formulate the Norm of Racial Equality as "don't be racist", and we realize that the Ideology of

Personalism is widespread, we begin to see how it is that one who accepts the Norm of Racial Equality might also assent to the items on the Racial Resentment test: as long as they don't classify those items as racist, they will take themselves to be adhering to the Norm of Racial Equality.

It is also worth noting that the Norm of Racial Equality is not, and has never been, in force for *everyone* in the United States. Mendelberg speaks of the Norm of Racial Equality as either being in operation or not: it was not in force for the United States in 1900, and it was in force for the United States in 1988. And for her purposes this makes sense. The voters that are her focus are those who politicians seek to sway via subtle racist manipulations—people with high levels of Racial Resentment, but who nonetheless adhere to the Norm of Racial Equality. But it is of course an oversimplification, for the norm may be in force for one group and not for another. Some Americans, those who identify as white supremacists, think it is perfectly acceptable to be openly racist; others do not.

Valentino et. al.'s 2016 paper shows that explicit racial appeals do not always fall foul of the Norm of Racial Equality in such a way as to block their influence on racially resentful voters. This suggests that *either* the norm is not in force, or it is no longer right to take it to preclude all explicit racial appeals. The latter is possible if explicit racial appeals are now seen as compatible with not being a racist. My suggestion in this paper is that this is now the case, at least for some explicit racial appeals. In the following sections, I sketch what seems to be one common technique for making an explicit racial statement, indeed often an explicitly racist statement, without being seen as racist: the figleaf. First, however, we need some more background in place on changes of attitudes and norms.

1.3 Changes in Attitudes, Norms, Permissibility

Rae Langton (2012) and Mary Kate McGowan (2012) tell a compelling story of how certain sorts of shifts in acceptability can take place. They begin from the way that what Lewis (1979) calls "conversational score" can shift. A crucial

notion here is that of accommodation: if a speaker says something which, for example, carries a presupposition, then—provided nobody objects—that presupposition is taken on board. When, for example, I say "my husband might be surprised by the thought that all feminists hate men", this introduces the presupposition that I have a husband, which can now be taken for granted for the rest of the conversation.

McGowan draws attention to the ubiquity of changes in conversational acceptability, arguing that every utterance changes—at least in a small way—what is acceptable for that conversation. Some of these changes will be small ones, like the need to take into account what was previously uttered. But others will be much larger—like a shift in the acceptability of racist utterances. Langton tells a similar story about conversational accommodation (though with a greater focus on the role of authority), arguing that that this can then bring about psychological changes in speakers, causing them to have attitudes and emotions that are appropriate to the racism now being taken for granted. Langton's and McGowan's discussions are especially focused on the way that openly racist utterances effect significant changes to standards of conversational acceptability. If these are made and not challenged, they maintain, the conversational score shifts so as to accommodate them as acceptable.

McGowan also notes that racist *behavior* may become permissible due to utterances that are made. And of course, this may lead to quite devastating consequences. Lynne Tirrell (2012), for example, has shown how the legitimation of hateful speech helped to give rise to the Rwandan genocide.

As so far told, however, this story is incomplete. To see this, consider what happens when a white supremacist makes an openly racist utterance—the sort of case that is McGowan's focus. If she is talking to another white supremacist, the remark will not be objected to. But this won't lead to any *change* in attitudes. The meeting of minds of two people with repugnant sentiments is obviously no good thing, but there is no reason to think that it will move others in a more racist direction. Now, consider what happens when a white supremacist makes

an openly racist remark to someone who adheres to the Norm of Racial Equality. Langton and McGowan suggest that if nobody objects, the racist presuppositions will generally be taken on board and begin to affect both the psychological states of conversational participants and permissibility facts more broadly.

But how likely is it that nobody will object? To make this question vivid, let's turn to the actual example McGowan uses (2012:121).

> Imagine that an African American man boards a public bus on which all the other passengers are white. Unhappy with the newcomer, an elderly white man turns to the African American man and says, "Just so you know, because I realize that your kind are not very bright, we don't like niggers around here… boy. So, go back to Africa… so you can stop killing each other… and do the world a favor!

It may well be that people will not *openly* object: confrontation is difficult, and people try to avoid it, even more so when race is at issue. Mendelberg's (2001, 2008 a, b) view was that open mention of race can disarm what would otherwise be an effective implicit appeal/covert dogwhistle: adherents to the Norm of Racial Equality self-monitor, and will reject what they cannot avoid seeing as racist. As noted, Valentino et. al.'s (2016) work has complicated this picture, meaning that open mentions of race may not block the effectiveness of an appeal. However, it does not by any means follow from their work that people will be untroubled by aggressive hate speech containing a taboo racial epithet, directed at an elderly man.[5] If, as I have hypothesized, even a thin norm of Racial Egalitarianism is still in force, *this* will be seen as unacceptably racist. It won't simply be seamlessly assimilated, but instead—in some way—rejected by those who adhere to this norm, even if they are racially resentful. This rejection might consist of mental distancing, changing the topic, or ending the conversation as

[5] Their research is on a different topic—explicit VS implicit racial political appeals, and their effects on the correlation between racial resentment and policy preference. They do not discuss an aggressive utterance of this sort (which is not in any way a political appeal), nor are they concerned with conversational norms and accommodation.

quickly as possible. But an adherent to the Norm is very unlikely to smoothly assimilate the racist assumptions, once they are seen as clearly racist. How, then, do norms shift? Is there something which allows the sort of assimilation that Langton and McGowan posit?

What is missing from their picture, it seems to me, is recognition of a further conversational phenomenon: the figleaf. A figleaf gives an openly racist utterance just enough cover that an adherent to the Norm of Racial Equality can reassure themself of the speaker's, and their own, non-racism. This is necessary to make an openly racist utterance seem like something that a non-racist might conceivably say and therefore crucial to the sort of conversational accommodation that Langton and McGowan draw our attention to. Without it, the psychology of speakers and the permissibility facts will not change. It is therefore vitally important for us to attend to figleaves and how they work in conversation[6].

2. Figleaves

A racial figleaf is an utterance made in addition to one that would otherwise be seen as racist. Unlike in the case of an implicit appeal/covert dogwhistle, race *has* been explicitly mentioned. The figleaf provides cover for what would otherwise have too much potential to be labeled as racist. Sometimes the figleaf is uttered at the same time as the racist utterance. However, as we'll see, figleaves can sometimes be provided as part of another, later conversation (they can even be provided in advance). The idea is that the figleaf offers some way of avoiding a confrontation with the possibility that something racist is going on. How well this works varies a great deal from context to context and audience to audience.

[6] I consider figleaves to be a friendly addition to McGowan's and Langton's accounts, which is wholly compatible with what they say.

This paper begins the project of exploring kinds of figleaves and their function. We will name some of the most common forms of figleaves, but we will also discuss some that are more complex and less easily labelled.

2.1 Synchronic Figleaves

A synchronic figleaf is one provided at roughly the same time as the utterance for which it is a figleaf. Probably the most easily recognizable figleaf is the classic "I'm not a racist but…", followed by something explicitly racial and quite possibly explicitly racist. Van Dijk (1993: 102-103) refers to this as an "apparent denial" of racism, and Hill (2008: 120) actually treats its intelligibility as a test of overt racism, noting that it only makes sense to use this phrase alongside something overtly racist.[7] We'll call it a Denial Figleaf.

The classic Denial figleaf attempts to dodge accusations of racism by simply asserting that they are not true. It is one of the most straightforward and crude of figleaves. Figleaves like this one are so well-known that there are entire blogs devoted to mocking them[8]. They are well known around the world. To take just one example, Van Dijk (82), writing in 1993, cites Jean-Marie LePen, then-leader of the Front National:

> (3) We are neither racist nor xenophobic. Our aim is only that, quite naturally, there be a hierarchy, because we are dealing with France, and France is the Country of the French.

A closely related figleaf is the Friendship Assertion figleaf, which often accompanies it. Its classic form is (2)

---

[7] I'm actually not so sure about this. I think it does make sense to say "I'm not racist, but I dislike Obama's foreign policy. While this may arouse suspicion that the *speaker* is racist—why else do they feel the need to assert their non-racism?—I don't think it indicates that dislike of Obama's foreign policy is a racist view, or that those who dislike Obama's foreign policy must be racist.
[8] http://imnotracistbut.tumblr.com; https://twitter.com/imnotsexistbut.

(2) Some of my best friends are black, but… [racist utterance].

I take another form of this to be to be the assertion of a fondness for the group attacked— as in "I've always had a great relationship with the blacks".[9] They are easily recognizable as weak attempts to deflect accusations of racism, and are generally not given much credence. For this reason, they are generally not very effective.[10] This is why more complex synchronic figleaves are also used, as discussed below.

I have defined figleaves in terms of their functions—they are additional utterances that serve to undermine the claim that an apparently racist utterance is racist. A Simple Denial figleaf will very often fail to do this. Technically, then, it will in these cases be a merely attempted figleaf.

2.1.1 Trump on Mexicans

Donald Trump's remarks about Mexicans include a synchronic figleaf:

> When Mexico sends its people, they're not sending their best. They're not sending you. They're not sending you. They're sending people that have lots of problems, and they're bringing those problems with [them]…They're rapists. And some, I assume, are good people.[11]

---

[9]http://www.huffingtonpost.com/2011/04/29/donald-trump-blacks-lawsuit_n_855553.html

[10] This, of course, varies substantially from subculture to subculture. It is also worth noting that there are generational aspects to this—older people may find it much more natural to insert Simple Denial or Friendship assertion figleaves, and may do so out of an abundance of caution. For example, non-racist members of certain subcultures/generations might quite sincerely utter "I'm not a racist, but..." before any comment remotely related to black people. It is entirely possible that one who asserts this is not in the slightest motivated by racism. A rather tragic consequence, however, will be that someone from outside the speaker's subculture might well attribute racism on the basis of the figleaf (where the comment on its own would not have triggered this attribution). I thank Dan Egonsson for pressing me on this point.

[11] https://www.washingtonpost.com/news/post-politics/wp/2015/06/16/full-text-donald-trump-announces-a-presidential-bid/

The first thing to note is that the quotation does not explicitly claim that all Mexicans are rapists. Instead, it explicitly associates Mexican immigrants with rapists, while at the same time putting in place figleaves. It focuses not on *Mexicans*, but on *the Mexicans who are sent.* This allows for both Trump and his supporters to insist that they are not prejudiced against *Mexicans.* Instead, they have much more specific negative beliefs about some Mexicans. And there's a further figleaf at the end: "some, I assume, are good people". This is a caveat added on, to allow a denial that the speaker is making sweeping generalisations even about those Mexicans who are "sent".

The addition of the figleaves to what would otherwise be very clearly racist generalisations leaves us with two claims:

> (4) They (the Mexicans who come to the US) are rapists.
> (5) Some of them (the Mexicans who come to the US) are, I assume, good people.

(4) is a generic claim about Mexicans who come to the US. Generic claims have notoriously slippery and confusing and controversial truth conditions, and those about social groups have been recently argued to have a crucial role in fomenting and perpetuating social prejudice.[12] But it's worth calling attention to two important (and widely accepted) facts about generics. First, they are not universal generalisations, but can be true even if there are exception. ("Cats have four legs" is true even though there are three legged cats.) This means we can make sense of (4) and (5) being both true without taking it that some rapists are good people. Second, they are nonetheless widely misunderstood as universal generalisations. (We see this every time a generalization about, say, the relative incomes of black and white people is met with the objection that there are some rich black people.)

---

[12] See, for example, Leslie forthcoming; Rhodes, Leslie and Tworek 2012; Wodak, Leslie and Rhodes 2015; Haslanger 2011. But for opposing views see Sterken 2015a and 2015b; Saul forthcoming.

This leaves the audience with an interpretation that can, on a very weak version of the Norm of Racial Equality, be understood as *not racist.* Because the claim is not about all Mexicans, and because there's an explicit recognition that even some of those who come to the US are (or are assumed to be[13]) good people, those who feel drawn to somehow associate Mexicans and rape can nod along while not having to see themselves as racist.

And this figleaf serves its purpose. Trump's defenders cite these points in order to argue that his comments about Mexicans were not racist.

> I think Trump is attacking them based on their actions, not their ethnicity. He is addressing the illegal immigrant group, not the race group they belong to. (https://www.quora.com/Did-Trump-really-make-racist-comments)

> I didn't hear him say anything racist against any race. What I did hear him say is, "Illegal Mexicans bring drugs, crime, and are rapists, but I'm sure some are good people." Seriously, whats racist about that? (Dirk, https://answers.yahoo.com/question/index?qid=20150728210521AAWJQfa)

> Trump is not racist…Trump is not against all mexicans just the illegals. (Julius Granstrom, https://twitter.com/juliusgranstrom/status/675231238366625792)

2.1.2 Glenn Beck on Muslims

Saba Fatima (2013: 341) discusses an utterance of Glenn Beck's that is also a synchronic figleaf. Beck is interviewing Muslim congressman Keith Ellison.

---

[13] It's worth noting that the addition of 'I assume' weakens the claim, by suggesting that Trump is giving them the benefit of the doubt without sufficient evidence. Still, the giving of this benefit of the doubt can be read as an indication of non-racism by one with a sufficiently narrow understanding of racism.

> OK. No offense, and I know Muslims. I like Muslims. I've been to mosques. I really don't believe that Islam is a religion of evil. I—you know, I think it's being hijacked, quite frankly. With that being said, you are a Democrat. You are saying, "Let's cut and run." And I have to tell you, I have been nervous about this interview with you, because what I feel like saying is, "Sir, prove to me that you are not working with our enemies." And I know you're not. I'm not accusing you of being an enemy, but that's the way I feel, and I think a lot of Americans will feel that way.

A part of Beck's utterance, (6), would seem quite clearly Islamophobic if uttered on its own to Ellison.

> (6) "Sir, prove to me that you are not working with our enemies."

One way that Beck avoids this is by using a group-based version of the Friendship Assertion Figleaf, (7).

> (7) I like Muslims.

But that's not all that he does. Crucially, he doesn't utter (6) on its own. He imbeds it in a rumination about what he *feels like* saying, mentioning it rather than using it.

> (6*) And I have to tell you, I have been nervous about this interview with you, because what I feel like saying is, "Sir, prove to me that you are not working with our enemies."

This we will call a Mention Figleaf. It allows Beck (and his supporters) to truthfully insist that he did not actually demand that Ellison prove that he is not working with the enemy. He avoids this speech act by mentioning rather than using the words that would, if uttered on their own, constitute the act of making such a demand.

And then he makes a further move. He continues with (8).

> (8) And I know you're not. I'm not accusing you of being an enemy, but that's the way I feel, and I think a lot of Americans will feel that way.

Here he explicitly states that he *knows* Ellison is not working with the enemy, which makes it rather puzzling that he also admits to feeling like saying "prove to me that you are not working with the enemy". Creating such puzzles is often crucial to the working of a figleaf, something we'll discuss more later in the paper. Finally, Beck finishes with an explicit denial that his utterance is an accusation. We see, then, at least three kinds of figleaves at work in Beck's utterance: Friendship Assertion, Mention, and a more complex further move with (8).

2.2 Diachronic Figleaf

A diachronic figleaf is one applied substantially later than the problematic utterance. Sometimes this is because attention has been drawn to the original utterance, and a response is demanded. Once more, the most obvious and crude versions are Denial and Friendship Assertion. Here we have Trump, being interviewed after making several utterances that were widely taken to show anti-black racism, such tweeting as false statistics about black crime taken from a white supremacist website and an expression of support for the assault of a black protestor at one of his rallies.

> (9) I have great African-American friendships. I have just amazing relationships, and so many positive things have happened.

On its own, this might not be terribly effective—it is so very close to the classic claim of a black best friend. But Trump also used a much more sophisticated dischronic figleaf. This was an utterance designed to *demonstrate* a lack of prejudice, rather than merely declaring it as a denial figleaf. In addition to his group-based Friendship Assertion Figleaf, Trump weighed in on the topic of

Supreme Court Justice Antonin Scalia's recent comments about affirmative action. Scalia had recently said "There are those who contend that it does not benefit African-Americans to get them into the University of Texas where they do not do well, as opposed to having them go to a less-advanced school, a less -- a slower-track school where they do well,"[14] comments met with widespread outrage. Trump's comments on the controversy came as a surprise:

> (10) I thought it was very tough to the African-American community, actually… I don't like what he said. No, I don't like what he said. I heard him, I was like, 'Let me read it again' because I actually saw it in print, and I'm going -- I read a lot of stuff -- and I'm going, 'Whoa!'[15]

(10) uses a much more effective maneuver than the Denial or Friendship Affirmation. It criticizes someone else for their racism, thus allowing the speaker to take the moral high ground and demonstrate what appear to be some anti-racist convictions. Once (10) has been uttered, Trump supporters can defend him against accusations of racism by noting that he criticized Scalia.[16]

3. How Figleaves Work—And Don't Work

3.1 Inference-blocking

A racial figleaf is, generally speaking, an attempt to block an inference from the fact that the speaker has made an openly racist utterance R to a claim like (11):

---

[14] http://edition.cnn.com/2015/12/13/politics/donald-trump-antonin-scalia-affirmative-action/

[15] http://edition.cnn.com/2015/12/13/politics/donald-trump-antonin-scalia-affirmative-action/

[16] Although the central notion of figleaf is that of an utterance, it seems to me that not-utterances can function just like figleaves. For example, it is commonly believed that no member of a group can be prejudiced against their own group. A racist utterance R, then, uttered by a member of the group that R is about, will generally be puzzling to the audience. They will hesitate to infer that either the speaker or the utterance is racist. In such a case, I am tempted to say that a person may function as a kind of human figleaf, but this requires broadening the notion of a figleaf beyond what I discuss here.

(11) The speaker is racist.

Given the ideology of personalism, this blocking will also have the result that the utterance R itself is no longer seen as racist. Blocking these inferences has a tremendous felt importance in contexts where the Norm of Racial Equality is in force. These are contexts, I've argued above, where conversational participants feel bound by a requirement to *not be racist*—however they interpret that. If they find themselves inclined to agree with a speaker, they are likely to hesitate if the speaker seems to have explicit racist commitments. Hill argues, rightly, that the ideology of personalism makes speaker intention the nearly-exclusive focus of any discussions that take place regarding racist language, allowing a wide range of denials based on ideas like mis-speaking, carelessness, or a good heart. My claim here is that figleaves are an important mechanism often involved in denials of racism, due to their ability to block inferences to claims like (11).

A Denial figleaf attempts to do this in the most direct way, by simply asserting the denial of (11). The audience in such a case is confronted with an utterance R that sounds racist, accompanied by the assertion that the speaker is not racist. On its own, the utterance of R might license an inference to (11) fairly quickly. But the Denial Figleaf attempts to block this. How successful this is will depend on a number of things.

If the utterance seemed clearly to be the kind of thing that *only a racist would say*, then the inference to (11) is very strong, and the audience will probably doubt the figleaf instead. And this will often be the case. As Van Dijk notes, "denials of racism are the stock in trade of racist discourse" (81). Similarly, if there is a great deal of other information pointing to the speaker's racism, the audience will probably doubt the figleaf instead of (11). Further evidence that might be used to cast doubt on the figleaf is knowledge that assertions of non-racism are very common among racists, due to the Norm of Racial Equality. The Denial Figleaf will only succeed if none of these factors cause the audience to reject the figleaf, which explains why it is only rarely successful (on its own).

Other figleaves do this in less direct ways. The Friendship Affirmation figleaf works by way of an inference from a claim like "some of my best friends are black" to the denial of (11). This is meant to be based, it seems, on the (incorrect) thought that a racist would not have close black friends. The Mention figleaf works by imbedding the utterance that would have licensed the inference to (11) within quotation marks. This makes it somewhat trickier to infer (11). Obviously, it's not the case that everyone who mentions a racist utterance is a racist. In a case like Beck's, however, (11) will still seem a reasonable inference to many of us. Nonetheless, employing the Mention figleaf renders the inference to (11) debatable in a way that it would not have been without it. Beck's defenders can insist that he did not demand that Ellison prove that he was not working for the enemy.

One way that a figleaf can be effective is simply by creating a state of confusion on the part of the audience. Trump's audience, for example, might feel that his utterances seemed potentially racist but that his criticism of Scalia casts doubt on this. In order to block the inference to (11), they need not actually reach any conclusion about his racism. A state of confusion and uncertainty will suffice to block the condemnation that may seem mandatory under the Norm of Racial Equality if (11) is endorsed.

3.2 Shifting Permissibility

Due to the Norm of Racial Equality, politicians attempting to exploit racial resentments need to be able to deny that this is what they are doing. Of course, it is far easier to make a convincing denial if you have avoided mentioning race. This is a significant advantage of using an implicit appeal/covert dogwhistle. However, figleaves can be used to provide deniability even when one has been more explicit. Indeed, as we have seen, this deniability may come in the form of simply denying racism, as in a Denial Figleaf. However, the more subtle figleaves offer more possibilities. Glenn Beck, criticized for his interview with Ellison, has ample potential to deny any racism by pointing out that he was very explicit about not accusing Ellison of working with the enemy. Donald Trump can insist

that he is not racist, and point to evidence of his non-racism, like his criticism of Scalia's remarks on affirmative action.

A figleaf has a dramatic effect on a conversation. In most situations, openly racist utterances create substantial discomfort. It is enormously difficult, socially, to accuse someone else of racism. But, for one who subscribes to the Norm of Racial Equality, it is also clear that explicit racism is not acceptable. A figleaf provides a way out of this massively disquieting impasse. If a figleaf has been uttered, there is room for doubt about the racism of the utterance, which removes the otherwise uncomfortably present obligation to object to racism.

And this is what potentially shifts the boundaries of the permissible. If nobody objects to a racist move, McGowan argues, the racism becomes acceptable. But, as we saw earlier, many explicitly racist utterances will *not* normally be smoothly assimilated where the Norm of Racial Equality is in force. A figleaf alters this dynamic. An effective figleaf allows explicitly racist utterances to be made, without objection. This means that the process of conversational accommodation is able to function in its normal smooth manner, adapting to the reality that R, the racist utterance, has been made and not met with any objections.

Figleaves (when they work) have the effect of defusing worries about racism. Once a figleaf has been effectively deployed, standards for what one can say without being racist shift. And this is powerfully worrying. If the audience accepts that the figleaf blocks the concern about racism arising from the utterance of racist sentence R, then R becomes s seen as something one can say without being racist. And this will make it far easier to say R, and even to do so without figleaves. Now we have our answer to how the boundaries of the permissible can shift. Among adherents to the Norm of Racial Equality, the crucial thing is to reject what is obviously racist. This allows an adherent to believe that they are following the very thin norm "Don't be racist!" And what counts as obviously racist can, and does, change. Pair something obviously racist with an effective figleaf enough times, and its racism is no longer obvious. At

that point, the figleaf may well begin to drop off.

A further effect is on how *other* utterances are perceived. As R+Figleaf makes its way into our discourse, slightly less racist utterances than R become unshocking. So, for example, Donald Trump's call to ban Muslims (accompanied by the figleaf "until our country's representatives can figure out what is going on"[17]) was initially shocking. As it was replayed on the news over and over, it became less shocking. And, crucially, the only slightly less racist call to ban all Syrians came to be seen as the moderate position in the Republican primary.

4. Problems/Complexities

4.1 How Effective Are Figleaves?

While it may sometimes happen that a figleaf is 100% effective and convincing for all audiences, this will be rare. Audiences will differ in the extent to which they accept a figleaf as casting into doubt the racism of an utterance. First, most obviously, the group targeted by the utterance is far less likely to accept the figleaf. There are surely very, very few Mexicans who doubt Donald Trump's bigotry toward them. But the reception of Trump's comments also shows us other complexities. While he has many admirers who deny that he is racist, many others have condemned his utterances as racist. The figleaves worked well with one group, but not with another. It is overly simple, then, to talk about a figleaf simply working or not working. Even a highly effective figleaf will be effective *with a particular group* and ineffective with others. Shifts in the boundaries of the permissible will also be circumscribed in this way: what becomes permissible within one community will not be permissible in another.

[17] http://www.npr.org/2015/12/08/458875362/trump-on-his-plan-to-ban-muslims-not-politically-correct-but-i-don-t-care.

4.2 Confused Figleaves

Self-knowledge is a difficult thing, and self-knowledge about racial attitudes is especially difficult, given widespread racial resentment combined with a norm demanding that one not be racist. Moreover, people have conflicting attitudes – famously, explicit commitments to anti-racism may coexist with implicit racist biases (see e.g. Brownstein 2016). This means that utterers of figleaves may not be deliberately engaging in manipulation. They may genuinely believe all the parts of their figleaf, and they may genuinely believe that the figleaf makes their utterance non-racist, or even that it demonstrates their lack of racism. One effect of this will be a blocking of self-understanding. Making a racist utterance, and having this be noticed and remarked on, can be a valuable turning point, which allows one to learn and change. Figleaves may block this from happening.

It may also be useful to think about figleaves at the level of belief, rather than utterance. The Norm of Racial Equality is not just a norm that causes one to worry about others considering one to be racist. It is also one that makes people not want to see themselves as racist. After saying (or even thinking) something racist, many people will find themselves worried—at least briefly—that they might be racist. At times like this, it is almost irresistible to seek reassurance by reaching for evidence that one is not racist. One kind of evidence can be a figleaf at the level of belief. One might think to oneself one of the obvious figleaves: *but I have a black best friend;* or *but I'm not racist.* Or, perhaps, one might move on to other topics that allow one to demonstrate one's lack of racism—quickly following a racist thought with a condemnation of someone else's racism.

4.3 Figleaves and Intention

It is important to note that a figleaf is defined in terms of its function, not the intention behind it. A figleaf is an utterance made in addition to an explicitly racist one, which provides cover by introducing doubt about the racism of the utterer, and therefore of the utterance. Sometimes, as with a racist but clever politician, this is intentional. However, as noted above, sometimes it is due to

confusion and concern over one's own possible racism—in a case like this it seems wrong to describe it as intentional.

I have mostly focused thus far on nefarious, deliberately designed figleaves, crafted to convince an audience that one's racist utterances are not racist. But I think it is a mistake to focus too much on intention. What matters most about figleaves is their *effects.*

It is in fact actually possible that Donald Trump thinks both that the Mexicans come to the US are generally rapists and murderers, and that some of them are good people. Certainly, a person who believed this *might* make the same utterances about Mexicans that Trump made, with no effort to conceal racism from either himself or others. He might simply believe both of these things. There need be no deliberate effort to manipulate. Importantly, however, this would make no difference to the classification of "some of them are good people" as a figleaf. This utterance still has the effect (for some audiences, anyway) of blocking an inference to the claim that the speaker is racist. And this is what matters for the classification as a figleaf.

To focus attention on the intentions behind a figleaf is to buy into the folk theory of racism that gives a central role to the Ideology of Personalism. The only difference is that the focus now is on the intention behind the figleaf rather than on the intention behind the apparently racist utterance. Our attention will still be, wrongly, focused on attempting to discern the state of mind of the speaker, and there will always be ways to doubt our judgments about the intention behind a figleaf.

But even where it is clear that a figleaf is an intentional manipulation, I think it is a mistake to engage in much discussion over intentions. Why? Because, as hinted earlier and argued below, what really matters about figleaves is their

ability to change what we see as permissible discourse, in extremely pernicious ways. And intention has nothing whatsoever to do with that.[18]

4.4 The Importance of Effects

A figleaf provides cover for an utterance that would otherwise be seen as clearly racist. This may in some instances be a deliberate manipulation of the audience by a fully explicit racist. It may in other instances be a back-pedalling by a genuinely conflicted person. It may even occasionally be an attempt to correct a genuine instance of mis-speaking. However, in all of these cases, a successful figleaf means that the inference from this utterance to the thought that the speaker is racist is blocked. An utterance of a sentence S that would have been clearly racist without the figleaf now comes to be seen as the sort of thing a non-racist might say. And since intentions and beliefs of the speaker are the most important thing according to the Ideology of Personalism, this means that S comes to seem not-racist, or at least not-clearly-racist. Crucially, this effect takes place no matter what intentions and beliefs lie behind the figleaf.

If figleaves do work in the way that I've suggested, leading us to change our views on whether an utterance is clearly racist or not, and facilitating the spread of racist speech, we should be very worried about their further effects. Lynne Tirrell's (2012) and David Livingstone Smith's (2012) work devastatingly demonstrates the ways that hate speech can lead to and be a part of genocidal violence. And they are not alone in this. Indeed, the United Nations condemns hate speech. And crucially, the UN's condemnation focuses not on the speech's intentions, but almost entirely on its effects— on speech that incites hatred:

---

[18] Defining figleaves in terms of inference-blocking effects has another consequence. We have so far looked at cases where the inference being blocked is one to a plausibly true belief. However, it will still be a figleaf if the belief being blocked is a false one. For example, consider the case of a feminist wanting to block the inference to the belief that she hates men. She might say "I don't hate men, but I think society is structured in a way that unfairly disadvantages women." "I don't hate men" is a figleaf to block the inference to the belief that the speaker hates men. The fact that she does not in fact hate men does prevent this from being a figleaf.

> all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin

Let's look again at an utterance-figleaf combination, and think about what the figleaf does and doesn't do.

> When Mexico sends its people, they're not sending their best. They're not sending you. They're not sending you. They're sending people that have lots of problems, and they're bringing those problems with us…They're rapists. And some, I assume, are good people.[19]

Without the figleaf, this is a clear case of incitement to hatred against Mexicans. You cannot call a group 'rapists' without inciting hatred against them. Do the figleaves mitigate this? Well, they allow for the possibility of good Mexicans who stayed in Mexico; and of an occasional good Mexican in the US. But if you have come to believe that, other than these exceptions, Mexicans are generally rapists, group-based hatred has clearly been incited. The figleaves do nothing to mitigate the hatred against the group. Indeed, by making it more socially acceptable— by calling into question the inference to the speaker's racism—they may increase the effectiveness of the incitement.

And this is why intention does not matter very much in the big picture, at least when we are discussing the public utterances of public figures, especially politicians. What is most worrying about these figleaves is their ability to make otherwise racist utterances seem acceptable, shifting our norms in such a way that increasingly explicit expressions of racial hatred become permissible. This

[19] https://www.washingtonpost.com/news/post-politics/wp/2015/06/16/full-text-donald-trump-announces-a-presidential-bid/

facilitates the incitement of racial hatred, which is and deserves to be a matter of the very greatest concern.[20]

5. Conclusion

Mechanisms that change our norms about racist utterances are vitally important. Those methods that allow previously unacceptable utterances to become acceptable can have devastating effects in the world. Right now, very dramatically, our social world is being altered by these methods. We need to notice it, and find a way to fight it effectively.

[20] It is far more appropriate to focus on intentions when considering utterances that do not have significant potential to reshape the norms of our world. In particular, when we are considering the utterances of those with whom we have personal relationships (e.g. family members), it may be especially important to focus on intentions. However, the focus on effects will remain important for utterances that do have significant potential to shape our social world. This will include not just public figures but others who have the ability to influence significant numbers of other people—teachers, journalists, lawyers, and so on.

Brownstein, Michael, "Implicit Bias", The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/implicit-bias/>.

Fatima, Saba. 2013. “Muslim-American Scripts”, *Hypatia* 28:2, 341-359.

Haslanger, Sally. 2011. “Ideology, Generics, and Common Ground” in Witt, Charlotte (ed.), *Feminist Metaphysics*. Springer Verlag 179-207.

Hill, Jane. 2008. *The Everyday Language of White Racism.* Chichester: Wiley-Blackwell.

Langton, Rae. 2012. “Beyond Belief: Pragmatics in Hate Speech and Pornography”, in Maitra, I. and McGowan, M.K., *Speech and Harm: Controversies Over Free Speech*, Oxford: Oxford University Press, 72-93.

Leslie, Sarah-Jane. Forthcoming. “The Original Sin of Cognition: Fear, Prejudice, and Generalization.” *Journal of Philosophy.*

Lewis, David. 1979. “Scorekeeping in a Language Game”, *Journal of Philosophical Logic* 8:1, 339-359.

McGowan, Mary Kate. 2012. “On ‘Whites Only’ Signs and Racist Hate Speech: Verbal Acts of Racial Discrimination”, in Maitra, I. and McGowan, M.K., *Speech and Harm: Controversies Over Free Speech*, Oxford: Oxford University Press, 121-147.

Mendelberg, Tali. 2008a. "Racial priming revived." *Perspectives on Politics* 6, 1: 109-123.

Mendelberg, Tali. 2008b. "Racial priming: Issues in research design and interpretation." *Perspectives on Politics* 6, 1: 135-140.

Rhodes, Marjorie; Leslie, Sarah-Jane; and Tworek, Christina. 2012. “Cultural Transmission of Social Essentialism.” *PNAS* 109 (34):13526-13531.

Smith, David Livingstone. 2012 *Less Than Human: Why We Demean, Enslave and Exterminate Others.* St Martins Press.

Sterken, Rachel. 2015a. “Leslie on Generics”. *Philosophical Studies* 172 (9): 2493-2512.

Sterken, Rachel. 2015b. "Generics, Content and Cognitive Bias". *Analytic Philosophy* 56 (1): 75-93.

Tesler, M. and Sears, D. O. 2010. *Obama's Race: The 2008 Election and the Dream of a Post-Racial America.* (University of Chicago Press.)

Tirrell, Lynn. 2012. "Genocidal Language Games", ", in Maitra, I. and McGowan, M.K., *Speech and Harm: Controversies Over Free Speech*, Oxford: Oxford University Press, 174-221.

Van Dijk, Teun. 1993. *Elite Discourse and Racism.* Newbury Park: Sage.

Valentino, Nicholas; Neuner, Fabian; and Vandenbroek, L. Matthew. 2016. "The Changing Norms of Racial Political Rhetoric and the End of Racial Priming", paper delivered to *APSA*.

Wodak, Daniel; Leslie, Sarah-Jane; and Rhodes, Marjorie. 2015. "What a Loaded Generalization: Generics and Social Cognition". *Philosophy Compass* 10 (9): 625-634.