

Inquiry

An Interdisciplinary Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/sinq20>

The social life of prejudice

Renée Jorgensen

To cite this article: Renée Jorgensen (20 Oct 2023): The social life of prejudice, Inquiry, DOI: [10.1080/0020174X.2023.2269197](https://doi.org/10.1080/0020174X.2023.2269197)

To link to this article: <https://doi.org/10.1080/0020174X.2023.2269197>



Published online: 20 Oct 2023.



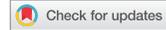
Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



The social life of prejudice

Renée Jorgensen

Department of Philosophy, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

This article considers a particular explanation (offered in Chapter 7 of Begby 2021) for the persistence of prejudicial stereotypes: that pluralistic ignorance can motivate individuals to act according to the roles they prescribe – even if no individual in a community either believes or endorses the stereotype – and moreover this can make it rational for subsequent generations to acquire prejudiced beliefs. I begin by surveying a few different ways that ‘vestigial social practices’ can persist despite being privately disavowed by most or all members of a community. Noting that many of them are transparently compatible with not believing that the persistent practice is appropriate, I argue that rational consideration of relevant alternative explanations precludes treating others’ behaviour as a kind of testimonial evidence for such prejudicial beliefs. But while it is doubtful that social dynamics provide grounds for rationally acquiring prejudice, it is likely that they explain actual acquisition of prejudice. So when evaluating whether a society is prejudiced, Begby is right that we must look beyond the private thoughts of its individual members. We should attend to the stabilising forces of social expectations, as well as how past prejudice shaped our material environment to reproduce stereotype-conforming social outcomes.

ARTICLE HISTORY Received 7 September 2023; Accepted 6 October 2023

KEYWORDS Stereotypes; social norms; rationality of prejudice

We often attribute to each other the beliefs and values that best explain our actions: seeing you drink the liquid in a cup, I might infer that you believed the cup contained gin, and that you like gin, and so accordingly drank the liquid from the cup. These sorts of inferences to the best explanation of behaviour do reasonably well at modelling the beliefs of an agent who navigates the world much as *homo economicus* does, deciding what to do based on what they think is most likely to be most first-order rewarding given their personal values.

But we don’t always navigate the world that way. We are deeply social, and frequently choose our actions with an eye toward coordinating with

CONTACT Renée Jorgensen  rjorgen@umich.edu  Department of Philosophy, University of Michigan, 2215 Angell Hall, 435 S. State St, Ann Arbor, MI 48109, USA

© 2023 Informa UK Limited, trading as Taylor & Francis Group

how others are likely act. This is often good: it makes it possible for us to cooperatively manage common pool resources, solve much more complicated social problems than we could manage individually, and to work together in a wide variety of domains.¹ The urge to coordinate with others motivates us to act as other expect us to, even when it conflicts with our narrow self-interest – often with pro-social results. But in some cases, this same impulse can lead us to collectively act in ways that are both harmful and come far apart from what any of us value or believe. Prentice and Miller (1993) documented a particularly clear instance:

Overdrinking. A group of undergraduates had a strong social norm of binge-drinking at parties. When interviewed individually, the majority of students reported that they personally preferred not to drink heavily, but thought (based on the others' public behaviour) that the majority of their friends considered binge drinking cool. So each student played along, drinking heavily in order to avoid being thought "uncool".

In this case, rather than motivating individuals to forgo self-interest for the sake of a greater social good, beliefs about how others expected them to behave functioned as a drag on progress, leading students to continue unhealthy drinking patterns they each privately dispreferred.

In Chapter 7 of *Prejudice: A study in non-ideal epistemology*, Endre Begby suggests that prejudiced beliefs – understood as high credence in sets of negatively charged generic stereotypes targeting a group of people – can likewise structure our social interactions even if no individual believes them. The core argument runs roughly as follows. Stereotypes set up mutually assumed social scripts, prescribing specific roles and interaction structures which we normatively expect each other to follow, and which enable us to anticipate and coordinate with others' behaviour. People play their roles when adequately incentivised, either by rewards or threatened sanctions. Begby calls particular attention to external negative sanctions as a stabilising force for prejudicial stereotypes: 'in violating script-based expectations, I risk being branded as "difficult", "awkward", or downright "clueless"' (Begby 2021, 117). To escape these (or more heavyweight) sanctions, it is prudentially rational to both avoid violating one's role expectations and avoid saying anything

¹Haslanger (2019, 11–13), drawing on McGeer (2007) and others, suggests that the social aspects of agency in fact go much deeper, providing the foundational scaffolding for human cognition and learning. Rather than considering individual agency as *prior* to social interaction and predicting how others will act, McGeer argues that when we attribute beliefs or predict behaviour we are partly 'giving ourselves over to the task of producing comprehensible patterns of well-behaved agency in ourselves and others' (2007, 149). On this picture, social coordination has a *regulative* and *shaping* role on individual agency.

that would bring one's commitment to fulfilling them into doubt. This leads to most people mostly 'assuming their roles in playing out these social scripts', whether or not they privately endorse them.

Imagine over time every individual in this society *privately* comes to disbelieve the stereotypes and disavow the associated social scripts. These changes in individuals' attitudes won't suffice to erase the prejudice from the society. As evidenced in *Overdrinking*, social norms can play a coordinating role when enough members of a community expect others to *expect them* to follow the norm – even if in fact few have the first-order attitude. Similarly, 'stereotypes can continue to govern the terms of our social interaction even though neither of us endorse their content', when 'both of us believe (or suspect) that the other endorses their content' (124). So it is possible for a society to be one in which interactions are organised by prejudicial social scripts without containing any individuals who believe the prejudicial stereotypes or endorse the corresponding scripts and social roles.

Begby notes that writ large, this is a mere theoretical possibility; in fact many people endorse both prejudicial stereotypes and the accompanying scripts. But he aims by highlighting the social aspect to offer a novel explanation of 'how stereotypes might come to have the social currency they do, and in particular, why they are so resistant to change' (132). He also suggests that the social aspect of prejudice often operates in local contexts as a kind of inertial deadweight against social progress, and offers the notion of '**perceived electability**' as a case study:

Let's say our country has never had a female head of state, partly as a result of a long history of institutionalized sexism. Along comes a female candidate who I believe is highly qualified, and indeed among the best candidates in the field. I would be happy to cast my vote for her. But at the same time, I believe that others will not vote for her, simply because she is a woman. Not wanting to waste my vote on an 'unelectable' candidate, I decide to go with someone I think has better chances of winning. But of course, a significant number of others might well be reasoning in the exact same way. (125)

The end result of this reasoning is that each of us *does* act as expected – voting for the suboptimal male candidate – *even though* we privately prefer the more qualified woman and do not privately believe that women are less fit to lead a country than men. We thus continue to be a society which will not elect women, because they are women, even though each individual disavows prejudiced attitudes or beliefs.

Begby leverages this to argue that individuals in the *next* generation could acquire prejudiced beliefs while neither themselves making any

rational errors, nor inheriting any irrationality from others. He draws on Susanna Siegel's (2017) discussion to note that collective behavioural patterns can provide something like social-level testimony for 'culturally normal beliefs'. Siegel illustrates how prejudice can be inherited with the example of Whit, a boy who grows up in a homogenous white community in the US and absorbs the racial attitudes and dispositions 'to be suspicious and distrustful of black men, to feel discomfort sharing public spaces' which are prevalent in his community. On Siegel's account, though absorbing culturally normal beliefs is not generally irrational, prejudiced beliefs formed in this way are ill-founded, inheriting the irrationality of the beliefs underwriting the behaviour of the individual members of the 'social-level testifier'. But in Begby's imagined cases, no one in the social-level testifier even *has* the belief to which their collective behaviour testifies. So if Whit were to absorb from them the prejudice that women aren't fit to lead a country, his 'attitude cannot be irrational in virtue of inheriting the irrationality of others' beliefs: if these people don't actually believe it, then they are *ipso facto* not irrational in believing it'. (129)

The interaction between individuals' prejudiced *beliefs* and a society's prejudicial *structures* and *scripts* is rich, and well worth careful attention from scholars interested in prejudice and non-ideal social epistemology more generally. I will in this article grant the hypothetical premise that the various members of the society have *indeed* cognitively moved on at the individual level, and are genuinely not themselves either credally or affectively invested in the perpetuation of the prejudicial norms and practices. But as many (including Haslanger 2019) have stressed, social roles and stereotypes are often embedded in webs of interdependent practices that jointly constitute the ways we navigate the world and make ourselves intelligible to each other *and ourselves*. Genuinely disavowing a central piece of this web – like gendered expectations concerning childcare – is a difficult feat, involving much more than lowering one's credence in the proposition that women will likely do more of the childcare tasks than men.

Though Begby's discussion is focused on explaining the resilience of prejudiced beliefs, this commentary will focus on the dynamics of the associated social practices, for two reasons. First, presumably part of the value of explaining the persistence of prejudice is to illuminate how to combat it. If so, then it is important to correctly identify the stabilising forces. Second, the existence of relevant alternative explanations for patterns of social behaviour plausibly undermines the rationality of treating the persistent practice as testimonial evidence in support of prejudiced

beliefs. If so, this casts doubt on the claim that Whit could *rationally* acquire prejudiced beliefs in the way outlined.

Pluralistic ignorance and the costs of coordination failure

There are unfortunately many cases where a social practice persists even though very few endorse it or its original justifying rationale. Let's call these 'vestigial social practices'. The undergraduate overdrinking norm has exactly the structure Begby attributes to prejudicial social scripts: the reason the practice persists is that each individual believes that the others *are* committed to the norm, and fears that they will apply external penalties (thinking them 'uncool') if she either fails to conform to the norm or even appears to not be committed to it – and each individual is ignorant that many of the others feel the same way. So, the perceived risk of social penalties motivates each individual to falsify the public record: not only conforming to the expectations, but refraining from asking questions that would make it visible to others that they are actually *not* personally committed to the norm.

This phenomenon – known as 'pluralistic ignorance' – is one reason a norm can continue to shape our practices even when no one endorses it: we collectively are simply not aware that it is unpopular, or at any rate are unaware that most others do not privately endorse enforcing the normative expectations it establishes. The good news about pluralistic ignorance is that it's possible to resolve by intervening on the social 'common ground': publicising the results of anonymous surveys, making public non-enforcement pledges, or even having highly influential cultural tastemakers visibly defect from the norm (Bicchieri and Mercier 2014, 64).² Were every undergraduate to be assured that no one would penalise failure to conform to the norm – because in fact no one is invested in its continuation, or because the coolest kids switched to mocktails – each would have no reason to continue overdrinking, and the norm would immediately destabilise. It is brittle, subject to defection cascades. When a social practice persists only because of pluralistic

²In some cases, participants are perfectly aware that few people privately endorse the norm, but still rationally expect to face second-order enforcement – penalties for either violating the norm or failing to penalize others' violations – and so the behavioural pattern remains stable until they can be credibly assured safety from penalization. Very likely Begby is right that many vestigial norms actually persist in part through pluralistic ignorance, precisely because we are actually still invested in many of the norms we publicly disavow – including many gender norms. For excellent discussion of how the interconnectedness of a wide array of social scripts can make it very difficult to in fact cognitively move on from socially embedded role-based expectations, see Bicchieri and McNally (2016).

ignorance, credibly learning that the norm will not be enforced is all it takes to move on. If this is the right model for the persistence of prejudice, then the way to intervene is to assure people that they will not be socially penalised for disclaiming or defying stereotypical roles and expectations. But not every social practice that coordinates behaviour is stabilised by threat of enforcement sanctions – even light ones, like branding violators as ‘difficult’ or ‘awkward’. Only the social practices that we normatively expect others to comply with, and in which we are invested, get that treatment.³

It is not impossible that universally (privately) disavowed negative stereotypes would persist due to pluralistic ignorance, but it isn’t straightforward to see how. For stereotyped individuals, there’s a high cost to conforming to the negative stereotype, and therefore some incentive to try to get away with violating the corresponding expectation. If not actually penalised, the visibility of violations would quickly erode both others’ empirical expectation that the stereotype will hold, and the normative expectation that others think it *ought* to. In these conditions, corresponding social practices would fade much like fashion fads and the concept of a beatnik.

Would violations be penalised? Begby’s imagined case stipulates that *no one* first-order believes the stereotype: no one personally expects that others *will* or should conform. They only believe that many *others* have these expectations. Under these conditions, we should expect people to penalise violations only if they in turn expect to be penalised for failing to *enforce* the norm. But if stereotypes take a descriptive form – purporting to merely describe regularities, rather than prescribe how people *ought to be* – there is nothing to ground this fear of second-order penalty. So pluralistic ignorance doesn’t look like a particularly good explanation for the persistence of structural social prejudice in the absence of any individual prejudice. But there are better models.

Vestigial social practices can also persist when they facilitate coordination around problems where failure to coordinate is particularly costly. Examples of this second type are plentiful, but driving conventions present a helpfully clear case. A *driving-side convention* solves an urgent coordination problem: we need to separate directions of traffic flow in order to avoid head-on collisions. It doesn’t matter which side we pick,

³For Bicchieri, this is the feature that distinguishes a *social norm* from a *convention*: norms are supported by an *empirical* expectation that others will act a certain way, together with a *normative* expectation that they believe we *ought* to behave in this way and will sanction departures. Conventions, by contrast, are supported solely by empirical expectations (Bicchieri 2006).

we just need to all drive on the *same* side of the road. But once we've picked a side, it can be *very* hard to change – even if everyone would prefer to switch. The Swedish driving side convention (discussed by Ullmann-Margalit 1977) illustrates this:

Left-side Driving. In Sweden circa 1960, the convention was to drive on the left. That was inconvenient, because all the neighbouring countries were right-side driving, so residents had to switch sides every time they crossed the border. Each person preferred to drive on the right, but as long as they expected the majority of other Swedes to actually drive on the left, it was rational for each to *also* drive on the left when in Sweden. And since they could expect others to reason similarly, even knowing that each other individual privately preferred to switch to right-side driving did not destabilize the left-side convention.⁴

Unlike *Overdrinking*, in *Left-side driving* credibly assuring everyone that no one would penalise others' violations would not destabilise the practice. Drivers need to coordinate their behaviour *in some way* to avoid head-on collisions, and left-side driving was the salient way. Until they could all coordinate around a specific alternative, it was prohibitively risky for any individual to violate the norm: they might collide with other drivers. The reasons that make it rational to continue driving on the side you expect others to drive on have nothing to do with norm *enforcement*; nor do they discourage people from voicing scepticism about the value of continuing the practice. Rather, they illustrate a second reason that a norm can continue to shape our practices even when no one endorses it: it is the salient solution to a coordination problem for which failure would be high-cost.

I think *Perceived electability* is better understood as this sort of case, rather than one in which we are responsive to anticipated penalties for violating a norm. Our anxiety about whether a given candidate is 'electable' makes no reference to the social penalties we might personally face in voting for a non-traditional candidate. Rather, it voices a concern about whether we can expect *enough others* in our party to do likewise to avoid a coordination failure: splitting the vote and losing the election. To destabilise the vestigial social practice in this sort of case, we need to establish an alternative coordination solution, *and* rationally believe that enough of us will switch to it at once to effectively

⁴As she explains, left-side driving remained the norm until changed by an official decree and months-long publicity campaign in 1972 (Ullmann-Margalit 1977, 88–89).

coordinate around the new norm and thus avoid all or most of the costs of coordination failure.

Plausibly quite a lot of social scripts – including some gendered scripts – have this form. Perhaps some men track sports news so that they can successfully participate in small talk on topics they know others are likely to raise with men; possibly a great deal of politeness behaviours more generally persist not because we are first-order invested in *those particular forms of expression*, or believe that these behaviours are in any way more natural or fitting than alternatives, but because it is important to have *some* mutually-intelligible ways of expressing social respect. When these scripts also encode prejudicial stereotypes, shifting them requires not just convincing people that the scripts are bad, but identifying and coordinating around specific alternative practices.

The social effects of material infrastructure

But there's another feature of *Leftside Driving* to which I want to call attention, which is especially important when considering how social prejudices persist. Even if Sweden had held a binding referendum and publicly voted to switch to right-side driving, there is another reason why social practices wouldn't shift overnight. A lot of physical infrastructure – the placement of road signs, lane indicators, merge lanes, traffic signals – had already been designed to support left-side driving. And because it is safest when driving on the left for the driver's seat to be on the right, not only were cars manufactured this way, but a host of car-related services (including drive through windows and toll booths) were oriented around the assumption that it would be. Until these material structures were changed to facilitate right-side driving instead, Swedes still had substantial practical reasons to keep driving on the left. This highlights a third reason that a vestigial practice can continue to shape our interactions even after everyone is 'cognitively ready to move on': our *material environment* is structured in ways that embed the practice into our lives. To eradicate this kind of practice, we have to intervene not only on individuals' attitudes and the common ground, but also on our material environment.

Arguably, most social prejudices – particularly sexist, racist, and ableist prejudices – are stabilised this way. Sexist beliefs might have initially motivated the choice to arrange economic and labour markets in ways that favour men. But well after the majority publicly disavows sexist attitudes and prejudices, if there are still material differences in the

affordances for men and women – asymmetries in parental leaves available, work hours expectations, or even gendered pay gaps – these constraints will make it (prudentially) rational for enlightened heterosexual couples to nevertheless instantiate a stereotypical gendered division of labour. This has nothing to do with anticipated risks of social sanction, and wouldn't discourage people from voicing scepticism about prejudiced gender beliefs. Asking critical probing questions about the underlying rationale for a practice of this kind won't budge it. It is much easier to destabilise a vestigial practice stabilised by fear of *enforcement* than one that persists because we shaped the material environment to support it.

It's perfectly possible – even common – for a social practice to be stabilised by all three forces (pluralistic ignorance, costs of coordination failure, and material structures), either simultaneously or in sequence. In the United States, racial residential segregation has been stabilised by a mix of these forces. Initially, widespread racial animus and viscerally prejudiced beliefs motivated legal codification and harsh enforcement of segregationist norms. Even after these were formally repealed, segregation was enforced informally through racially restrictive real-estate covenants that prohibited the sale of valuable properties to non-white purchasers, redlining practices that made mortgages unattainable for Black applicants, and social sanctions visited on employers, businesses, and individuals who disregarded segregationist social norms. Pluralistic ignorance played a role in how long these informal norms persisted: in several studies (as early as 1931), the overwhelming majority of white subjects interviewed reported that *they personally* were opposed to segregation, but even in 1968 they would not support anti-segregationist policies because they believed that *most others* were pro-segregation.⁵

Even after the majority of white Americans publicly opposed residential segregation, it persisted (and has even recently intensified).⁶ Neither repealing segregationist statutes, nor legally prohibiting discriminatory lending or sales was enough to reverse segregated residential patterns. This is in part because segregation was the explicit social order during the main initial phase of suburbanisation, and many aspects of municipal

⁵In an early study of this form, Katz and Allport (1931, 152–157) found this pattern among undergraduate fraternities. In analysing responses to a 1968 survey, O'Gorman (1975) found that whites who neither strongly favoured desegregation nor segregation supported whichever policy they believed the majority of whites supported, and 'in 1968 most white Americans grossly exaggerated the extent to which other whites supported racial segregation.'

⁶Medvedian, Gambhir, and Gailles (2021) found that 81% of metropolitan regions in the United States with more than 200,000 residents were more racially segregated in 2019 than they were in 1990.

and social infrastructure – including the locations of freeways, heights of overpasses, design and location of public transit, and residential zoning codes – were built up to support and reinforce its effects. It's also in part because of the compound economic consequences of past discriminatory residential policies. American suburbanisation was nearly complete by the time the 1968 Fair Housing Act prohibited future racial discrimination, and since then housing prices have soared while wages stagnated. The difference in generational wealth between families that were permitted (or given loans to enable them) to purchase property prior to 1960 – and so accrued equity as prices rose – and those that weren't is staggering.⁷ It also means that now facially neutral policies – like tax deductions for mortgage interest, or evaluating loan eligibility on financial assets alone – have effects that further entrench racial divides in home ownership and social mobility.⁸

Changing private attitudes about the desirability of integration, or the raw abilities or value of the labour performed by women and people of colour does not on its own make them more likely to have a high net worth, be chief executives, or own homes (let alone in diverse neighbourhoods). Nor would making it common knowledge that no one believes that they are intrinsically any less capable of performing highly compensated labour. While individual prejudiced attitudes played a causal role in setting up the social infrastructure, they are not necessary for the continuation of the patterns asserted by these stereotypes. We have structured the material environment in ways that make them highly probable social outcome, independent of individuals' attitudes. We'll return to this point in a moment.

Is it really rational to acquire culturally normal beliefs?

We've quickly glossed three reasons (pluralistic ignorance, coordination costs, and material infrastructure) why social practices can persist once they have emerged, *other than* that the truth of – or even widespread belief in – the prejudicial stereotypes initially used to rationalise the practice. In addition to shedding some light on the range of interventions necessary to root out structural prejudice, this fact has implications for the rationality of acquiring or absorbing beliefs from widespread cultural practices. Let's consider Whit again. According to Begby, Whit 'merely

⁷In the United States, median home prices increased at four times the rate of household incomes since 1960, while rents have increased at twice the rate.

⁸For discussion, see generally Rothstein (2017, Chpt 11) and Anderson (2010).

observes others going about their everyday business, reasonably infers that they hold certain beliefs, reckons that if they all believe it, there's a good chance that it's true, and thereby concludes that he should believe the same himself'. (129) He argues not only that Whit is rationally *permitted* to acquire prejudicial beliefs this way, but that were we to advise someone in Whit's position that they *not* do so, 'we would be counseling people to disregard what they have every reason to believe is evidence' (129). Given what we've observed about vestigial social practices, the argument here merits a closer look.

'Complying with stereotypes in one's outward behavior' (133) could mean many non-equivalent things. Most permissively, it could mean simply behaving in a way that is not *inconsistent* with the stereotype: not being a counterexample, not vocally arguing that the stereotype is misleading. Alternatively, it could mean doing things that are *best* explained by attributing to you belief in the predictive validity of the stereotype. Most narrowly, it could mean appearing to be a *confirming* instance – though it isn't clear anyone except the target of a stereotype can do this. Begby offers the following gloss when discussing *Perceived electability*:

we all mold our behavior to fit the standard social script according to which women lack essential leadership qualities, even though none of us fully endorse the underlying stereotype. In doing so, we are also facilitating the continued suppression of the very evidence that would actually demonstrably problematize the underlying stereotype, namely a record of female politicians competently performing the duties of public office. (125)

This suggests the more permissive interpretation, and implies that a broad range of behaviour – anything *consistent* with believing the stereotype – counts as complying with it. But is it true that *inference to a salient explanation* of others' behaviour in this capacious sense – especially of *omissive* behaviour (e.g. the absence of integrated classrooms, or failure to elect women to political leadership positions) – provides a rational basis for forming a *specific* belief like 'it is better to preserve segregation' or 'women lack leadership qualities'? Begby argues that 'socially accepted patterns of action and interaction can provide individuals with something like a non-overt form of testimony even in the absence of explicit assertion' (127). But overt testimony clearly identifies the proposition for which one is receiving evidence. Observations of others' behaviour is not only non-overt, it radically underdetermines what beliefs might motivate it.

To make this more concrete let's look at another of Begby's illustrations. It is presently a culturally normal belief that 'MSG is bad for you'. Though trumpeted by explicit testimony when I was a child, it has since been discredited but has not disappeared from the social world. Begby argues that

One cannot *not* respond to the widespread MSG-scare and the 'No MSG' signs in Asian restaurants as evidence, grounds for some inference or other. But if it is evidence, what is it evidence for? Which inference is warranted? Presumably that MSG is bad for you.

I'm not so sure. Suppose we grant that it is true that I cannot *not* respond to any fact as evidence for some proposition. I can nevertheless quickly irrationally over-reach my support, or misidentify the proposition for which it is support. Along with signs announcing the absence of MSG, I also frequently encounter signs and labels exclaiming 'Gluten-Free!', 'Not from Concentrate – ever!' and 'Contains Nuts'. It seems to me that someone who concluded from these labels that gluten, concentrate, and nuts are *probably bad for you* would be irrational; at most what these labels give them evidence for is just that *some consumer behaviour is responsive to that ingredient*.

So back to the case of prejudice. Whit observes few women in politics or at high-level executive positions, and high levels of racial residential segregation. He notes that these are stable patterns: they have been true for the whole history of his country. Is it rational for Whit to respond by treating these observations as evidence? Are they good grounds for his inference that most people think women lack essential leadership qualities, or that racial segregation is desirable, and if most people believe these things, they are probably true? Set aside this last step for a minute to focus on the rationality of the first inferential step. If – visiting Sweden in 1971 – Whit were to observe that most Swedes drive on the left side of the road, and have done for as long as they had cars, and then took this as evidence for the inference that Swedes believe left-side driving is more efficient or natural than right-side driving, we would *not* happily accept that his inference is rational. First, his evidence is deeply equivocal: many different background beliefs are compatible with driving on the left. Second, in leveraging the persistence of the practice to attribute this belief to the Swedes, Whit fails to consider relevant alternative explanations for the stability of their practice (e.g. that the drivers are merely following the laws, or aiming to avoid collisions) that do not involve believing that left-side driving is best. To reason

from others' observed behaviour to a particular best explanation, rationality requires not only that one consider relevant alternative explanations for the behaviour, but also that one has some reason to dismiss those alternatives.

We surveyed three alternate explanations for the persistence of social practices that embed racial and gendered stereotypes, compatible with it being false that most people believe the stereotypes. If it were quite difficult to learn that these are candidate explanations of the persistence of the social practice, *perhaps* Whit could be still counted as rational when, failing to consider them, he instead infers that the stereotypes are widely believed. But the suggestion that inequitable social patterns persist because of social structures and inequitable opportunity isn't just whispered in the dark corridors of the Ivory tower; it's properly publicised, broadcast in mainstream popular forums. And even if the technical terms are not widespread, the *concept* that norms can persist because of pluralistic ignorance or the costs of coordination failure is easy to grasp. If Whit considers these alternative explanations, then if he is rational, he would recognise that his evidence does not support inferring that the stereotypes are widely believed. If he fails to consider them, he violates a rational norm. Either way, *this* is not a path to rationally acquiring prejudiced beliefs.

At some places, though, Begby speaks in terms of 'enacting stereotypes that none of us actually endorse' (133), and points to how our behaviour *suppresses* the emergence of counterexamples that would unseat the prejudicial stereotype. Rather than merely setting Whit up for inferences like the ones above, perhaps the thought is that conforming behaviour generates *direct* stereotype-confirming evidence. If this is the view, the terrain is even trickier.

Some stereotypes make themselves true. That is, reinforced by the infrastructural legacy of historically endorsed prejudice, they continue to make the features of a stereotype non-accidentally correlate with each other. Men *are* more likely to engage in conversations about sports; women *are* unlikely to be elected to political office; Black Americans *are* less likely than white men to hold executive-suite positions in companies; American neighbourhoods *are* unlikely to be racially integrated. Whether it is rational to believe a stereotype is in these cases is a complex question, for reasons related to what Munton (2019) calls the 'Problem of Projectability'. The problem occurs when a statistical observation is projected onto a novel instance outside the conditions that underwrite the observed regularity. It could be *true* that in my

country less than one per cent of successful political leaders are women, and that fact could make it rational for me to have a high credence that the female candidate on the ballot is unlikely to be a successful leader – but that does not make it rational for me to believe that *even if she were to win the election* she would be unlikely to successfully lead.

What does this mean for Whit? He can rationally believe that under present social conditions, the female candidate is unlikely to be elected. But given the relevant possibility that the observed social patterns are the downstream result of social infrastructure and practices, he should be very cautious when drawing inferences that would project these patterns beyond the modal profile of the social practices. Inferring that they hold because of features *internal* to the individual instances is a paradigmatic move of this kind and is common in most prejudicial stereotypes. In other words, he would overreach his evidence dramatically were he to infer that women lack the qualities necessary for political leadership – or even that most people around him believe that they do.

Wrapping up

Where does all this leave us? Though I have emphasised points where I think the details of the account Begby has offered need to be refined, I think he is absolutely right to identify vestigial social practices as a principal reason for the stubborn persistence of prejudicial stereotypes. Shifting the focus from questions of individual people's attitudes to the scripts and structures of our shared social life is a move in the right direction. His attention to the material and structural forces that make it hard to dislodge stereotypes and prejudice is an advance over thinking that the sole difficulty lies in the head of individual prejudiced believers.

Even beyond that, he offers a plausible story for how individual-level prejudice would re-emerge in the next generation in a society in which every individual had *privately* rejected prejudiced attitudes and expectations, but which retained prejudiced vestigial social practices. Probably a character like Whit *would* reason in exactly the ways that Begby outlined. However, it is dubious that the beliefs absorbed in this way would be *rationally* acquired. Nor would we be recommending that anyone violate a rational norm, were we to urge people in Whit's position to refrain from using stable social practices as an inferential base for beliefs about e.g. women's aptitude for political leadership.

I'd like to take the opportunity in closing to tie together some broader themes from the book. Begby readily – and rightly – grants that even if it

would be rational to believe that the female candidate is unlikely to win the election, we would do something *bad* if we took that fact as justification to not vote for her. Doing this is *precisely* how our behaviour perpetuates current structural prejudice. This, I submit, is the feature of stereotypes that makes many of them so persistent. It also underscores the importance of being very careful about the purposes to which we put the kinds of generalisations that often underwrite stereotypes.

Whether I am relying on a generalisation in a way that falls afoul of the problem of projectability depends in part on *how* I am trying to use it. If I am merely trying to predict the next set of social outcomes, I can legitimately hold fixed the actual social conditions, and so stay ‘within bounds’, so to speak. But if I am trying to guide an intervention on those conditions, or attempting to match treatment to an individual’s intrinsic underlying merit – as is plausibly the task when using risk scores to decide whether to make someone eligible for early parole, or set an interest rate for repayment of a mortgage – it is not *just* that the ethical stakes of my decision matter (though they do), or could override epistemic norms (though they might). Rather, the fact that the shape of my inference applies the generalisation beyond its modal profile is a narrowly epistemic flaw in my reasoning: the norms of epistemic rationality do not permit this sort of inference. Begby is right that prejudice is persistent because it is bound up in our social practices, but wrong that this is because they make it rational to believe prejudicial stereotypes. Rather, it is because our social practices are supported by material infrastructure that both creates conditions under which it is difficult for counter-evidence to emerge, and is often ignored when, noticing the stubbornly vestigial patterns, we ask why they hold even after we no longer think that is how things *ought* to be arranged.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Anderson, Elizabeth. 2010. *Imperative of Integration*. Princeton: Princeton University Press.
- Begby, Endre. 2021. *Prejudice: A Study in non-Ideal Epistemology*. Oxford: Oxford University Press.
- Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.

- Bicchieri, Cristina, and Peter McNally. 2016. "Shrieking Sirens: Schemata, Scripts, and Social Norms. How Change Occurs." *Social Philosophy & Policy* 35 (1): 23–53.
- Bicchieri, Cristina, and Hugo Mercier. 2014. "Norms and Beliefs: How Change Occurs." In *The Complexity of Social Norms*, edited by Maria Xenitidou, and Bruce Edmonds, 37–54. Cham: Springer.
- Haslanger, Sally. 2019. "Cognition as a Social Skill." *Australasian Philosophical Review* 3 (1): 5–25.
- Katz, Daniel, and Floyd H. Allport. 1931. *Student Attitudes*. Syracuse/New York: Craftsmen Press.
- McGeer, Victoria. 2007. "The Regulative Dimension of Folk Psychology." In *Folk Psychology Re-Assessed*, edited by Daniel D. Hutto, and Matthew Ratcliffe, 137–156. Dordrecht: Springer.
- Medendian, Stephen, Samir Gambhir, and Arthur Gales. 2021. *Twenty-first Century Racial Residential Segregation in the United States*. Berkeley: University of California at Berkeley.
- Munton, Jessie. 2019. "Beyond Accuracy: Epistemic Flaws with Statistical Generalizations." *Philosophical Issues* 29 (1): 228–240.
- O’Gorman, Hubert J. 1975. "Pluralistic Ignorance and White Estimates of White Support for Racial Segregation." *The Public Opinion Quarterly* 39 (3): 313–330.
- Prentice, Deborah A, and Dale T Miller. 1993. "Pluralistic Ignorance and Alcohol use on Campus: Some Consequences of Misperceiving the Social Norm." *Journal of Personality and Social Psychology* 64 (2): 243–256.
- Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York: Liveright Publishing Co.
- Siegel, Susanna. 2017. *The Rationality of Perception*. Oxford: Oxford University Press.
- Ullmann-Margalit, E. 1977. *The Emergence of Norms*. Oxford: Oxford University Press.