Jessie Munton

Prejudice as the misattribution of salience[1]

(9902 words)

Abstract: What does it take to be prejudiced against a particular group? And is prejudice always epistemically problematic, or are there epistemically innocent forms of prejudice? In this paper, I argue that certain important forms of prejudice can be wholly constituted by the differential accessibility of certain pieces of information. These accessibility relations constitute a salience structure. A subject is prejudiced against a particular group when their salience structure is unduly organised around that category. This is significant because it reveals that prejudice does not require the presence of any explicit cognitive or emotive attitude, nor need it manifest in behaviour: it can be solely constituted by the *organisation* of information, where that information may be accurate and well-founded. Nonetheless, by giving an account of 'undue organisation' in epistemic terms, I show that this account is compatible with an understanding of prejudice as a negatively-valenced epistemic category.

Tell me where is fancy bred,
Or in the heart or in the head?
How begot, how nourished?
Reply reply.
(William Shakespeare *The Merchant of Venice* 3.2.63)

Introduction

What does it take for an individual to be *prejudiced* against a particular demographic group, to be racist, or sexist, for instance? There are three main ingredients that feature in answers to this question. According to many, prejudice requires a particular *cognitive attitude* - the endorsement, explicit or implicit, of a proposition of some kind - that the group in question are in some way inferior, for instance. Some accounts add an epistemic condition: that the proposition in question be false or irrational. The second element is a negative *feeling* towards the group in question, such as dislike or disgust. And the third ingredient is behaviour. Most accounts are hybrid, requiring some combination of those factors.[2]

In this paper, I offer a minimal account of what it takes to be prejudiced, which denies all three of the proposed conditions above. Prejudice need not manifest in behaviour, nor does it depend on a negative emotion, or the endorsement of any particular proposition about the group in question. Instead, prejudice can arise purely through the *organisation of information.* Information is organised both by an individual's mind and their broader social context into what I shall call a salience structure, understood as an ordering of information by accessibility. The central claim of this paper is that problematic salience structures, salience structures which are unduly organised around

[1] For feedback on earlier drafts of this paper I am indebted to Georgi Gardiner, Huw Green, Gabrielle Johnson, Susanna Siegel, and to audiences at Robinson College, Cambridge, a workshop on the ethics of belief at the University of Köln, the Bled epistemology conference, and an on-line workshop on salience hosted by Harvard University.
[2] See Levy (2016) for this threefold distinction between doxastic, behavioural, and affective standards when he employs when assessing whether implicit bias against other racial groups amounts to racism

demographic categories, are sufficient for prejudice, even when the information which falls within the ordering determined by the salience structure in question is not itself inaccurate, and regardless of whether that then manifests in behaviour, or is accompanied by negative affect.

This has some significant upshots: it offers us a minimal account of prejudice which explains a number of hard cases including microaggressions and prejudicial ignorance. In doing so, it identifies a common core to what otherwise appear to be a range of disparate phenomena which we intuitively recognize as prejudiced. And it meshes with recent work which seeks to move the focus of our understanding of prejudice away from a focus on individual intention or explicit belief.

In the first section of the paper I set out the philosophical context for the claim I am making. In the second section I motivate the account by discussing three phenomena: perseveration, selection and abduction, and argue that all are rooted in a default ordering of information. It is tempting to think of these cases as unified by problematic patterns of attention, but as I argue in section four, an analysis that settles at the level of attention fails to arrive at any common core to the phenomena of interest. Instead, we need to look at the salience structure which *drives* those patterns of attention. In section five I offer an account of what a salience structure is. In section six I set out what it takes for a salience structure to be prejudicial. In section seven I deal with some objections and draw out further implications of the account.

1. What does it take to be prejudiced?

What does it take to be prejudiced?[3] What does it take to be racist, for instance, or sexist, or homophobic, or prejudiced against conservatives or older people? One natural answer to this question requires the subject to hold explicit negative beliefs about the group in question, for instance that women are less able than men in some respect, or that people of a particular race are inferior. It requires, in Michelle Moody-Adams' phrase "a distinctive conception of the nature of reality" (2007, 91). It is very natural to add the requirement that such beliefs must be inaccurate, or irrational, or epistemically flawed in some other way (Appiah 1990).

Must one also dislike the group in question? Must prejudice be in the *heart* as well as the head? Some philosophers have proposed affective conditions, as a supplement or alternative to the cognitive. Jorge Garcia, for instance, argues that racisms requires individual *ill-will, "*a vicious kind of racially based disregard for the welfare of certain people*"*) (Garcia 1996, 6). Garcia's account, like Moody-Adam's, is also an instance of a *moralized* account: he argues that racism must involve a moral vice.

[3] Why ask about prejudice rather than some more specific form of it, such as racism or sexism or antisemitism or homophobia? One might wonder whether 'prejudice' *per se* even pick out a distinctive category with any significant features in common among its members. Take the more specific term racism: there is dispute over whether even that picks out a single, unified phenomenon. Lawrence Blum (2004) argues that ordinary usage of the term does not, contrary to, for instance, Joshua Glasgow, who argues that it is unified by disrespect (2009). There are in addition many subtler distinctions to be made here. Anthony Appiah, for instance, distinguishes between racism and racial prejudice, using the latter to pick out "the deformation of rationality in judgment that characterizes those whose racism is more than a theoretical attachment to certain propositions about race (1990, 8).

The reason for the generality of my target is that I am interested in a feature of our mental organization which supports a wide range of more specific forms of prejudice. Of course this limits the analysis - there will be important differences among these various forms of prejudice, which this account cannot capture. I encourage readers to bear these potential distinctions in mind as they read, and to be alive to what the breadth of the target lets us accomplish, and what it prevents us from doing successfully.

It can also be natural to impose a behavioural criterion: that sexism requires *oppression* of women, for instance (Cudd and Jones 2007), or that racism requires acting towards people of another race in a way that manifests that negative belief, or that has discriminatory consequences (Appiah 1990) or that engages in ethnic discrimination (Corlett 1998).

Much recent literature has been concerned with *implicit* bias: prejudicial attitudes that are revealed by behaviour, for instance on tests designed to detect patterns of association, but that may be at odds with the individual's explicitly avowed beliefs or behaviour (Greenwald, McGhee, and Schwartz 1998). The phenomenon of implicit bias raises (at least) two species of further questions. One is whether it amounts to a full-blown prejudicial attitude such as racism, particularly in individuals who endorse explicit anti-prejudicial commitments, for instance (Levy 2017). The second species of question concerns the psychological underpinnings of the attitude. What kind of a state is it? What psychological apparatus does it rely on? Are these implicit attitudes associations between concepts, or are they propositionally structured? (Mandelbaum 2016).[4]

These debates within the literature on implicit bias mirror two distinct but related questions that underlie much of the literature on prejudice more broadly. The first asks what conditions something must meet in order for it to constitute an instance of prejudice – it describes the phenomenon in *functional* terms. The second asks what (psychological) mechanisms realise those conditions. The two are linked in so far as some answers to the first question set requirements on the second (an explicit belief, for instance). With this framework in place, let me situate the account I offer here.

My primary aim is to offer an answer to the first question which in turn opens up some novel space for how we answer the second question. The organisation of information, described in terms of a salience structure, can be sufficient for prejudice. Prejudice need not be in the head or the heart, so much as in the filing cabinet.

Note that I am not claiming that a malfunctioning salience structure is *necessary* for prejudice. It is *one way* that prejudice can be constituted. And I am certainly not arguing that certain kinds of prejudice do not require beliefs, affective attitudes or a behavioural manifestation of some kind. Rather, the claim is that there could be some minimal cases of prejudice which can be constituted solely by a prejudicial salience structure. That structure is in addition likely to underlie discriminatory behaviour, inaccurate and irrational beliefs and to have a symbiotic relationship with a range of affective attitudes. Identifying the role of informational organization helps us better understand prejudice *across the board* and in particular those instances of it where it is constituted by that organization.

That answer to the first question has complex ramifications for the second question: on the one hand I will remain fairly neutral on the question of how we implement this organisational structure at a psychological level. In particular, I set no explicit requirements on the nature of the information that is organised. The claim is compatible with both propositional and associationist approaches to implicit bias, for instance. But I am committed to an architecture of the mind on which the

[4] Johnson (forthcoming) offers an account of bias that cuts through a number of these debates. She argues that characterising bias in terms of its functional role – described in terms of representational mental states that serve as inputs and outputs - lets us remain neutral on the nature of the mental states that bridge the gap between them, thereby allowing us to recognise the diversity of states and mechanisms that can fill that implementational role. In particular this lets us move away from the focus on the question of what representation an individual harbours, towards an interest in how those representations are manipulated. The account I offer here constitutes a move in a similar direction, since it focuses on the *accessibility* or organisation of information.

organisation of information is distinct from the information itself. The account I offer here is inconsistent, for instance, with a view on which the organisation of information in the mind simply reduces to a further set of beliefs, implicit or explicit, about what information is most important.

2. Salience

Consider, as an entry way into the account I want to offer, the attitude of familial love. Suppose that Esther is a generally average child who is much loved by her parents. Esther's parents hold a range of true beliefs about their child - in fact, we can specify that they are clear-eyed enough that almost all of their beliefs about her are true. They believe that she dislikes washing her hair, has a good range of funny faces, is sometimes unkind to her younger sister on purpose, enjoys colouring in, is frequently unreasonable and so forth. Their belief set covers the gamut of positive and negative beliefs. This is not to say that Esther's parents are impartial where their offspring are concerned, however. Far from it: when Esther's parents think about or discuss their child, they more often entertain or articulate the positive beliefs than the negative ones. When they "reach for" a fact about or feature of their child, it's as though the positive ones are easier to access. This isn't to say that they *cannot* access the negative when relevant, it just takes a little more prompting or effort. If they are having a conversation with one another about how Esther will fare when starting school, they can bring to mind readily enough the fact she doesn't like to share with other children, but sooner rather than later one of them will revert to reflecting on some more positive aspect of her character: how bold and exploratory she is in new situations.

Esther's parents are importantly biased about their child, but that bias need not take the form of *false beliefs* about her. Instead, it can be *constituted by* an ordering of their beliefs - some are more accessible than others, and that ordering in terms of accessibility is my focus of interest here. Emotions of familial affection are complex (Freud 1899), but part of what constitutes them is an accessibility ordering on the beliefs one holds about one's family members. That ordering is part of what it is to love someone.

One way of describing what is happening in the case of Esther's parents is in terms of *salience.* When they think about their daughter, positive information which they have about her is more salient than negative. But what exactly do I mean by *salience?*

Salience can be a property of objects, or of their properties. A salient stimulus stands out from others, drawing our attentional resources to it (Gaspelin and Luck 2018). As I use the term, a salient object or property may be external to the mind - the red postbox on the grey street corner, for instance - or it may be internal to the mind, a belief, or visual experience which demands attention.

Salient stimuli may be very low-level. A black cat seen in white snow will be perceptually salient: in virtue of the contrast between it and the background, most typical viewers will have their attention drawn to that stimulus. But far more complex and sophisticated stimuli can also be salient. Suppose you get on to a bus in the middle of the day, and see an adult wearing a large rubber chicken head over their face. That will be a salient stimulus, whether or not there is any low-level perceptual contrast between the colours of the head and the surrounding bus. It is salient because of it is unusual in that social context. Its peculiarity demands explanation: *why* is this individual dressed so oddly? That kind of salience relies not just on facts about low level processing common to different individuals, but also on culturally specific knowledge about normal behaviour on buses.

Whilst an individual is not able to determine at will what stimuli are salient to them at a given

moment, patterns of salience are shaped by social identities (Humphreys and Sui 2015) and goals (Lien et al. 2008; Kamke and Harris 2014) that an individual may cultivate over time. For instance, a long-term interest in a particular football team could make a jersey on display in a shop window salient. Whilst the individual cannot straightforwardly control the salience of the jersey at that moment, they may voluntarily engage in many behaviours (attending football matches, watching games on TV, discussing sport with their friends) which are partly responsible for the salience of the jersey.

The attribution of psychological salience is mediated by the neurotransmitter dopamine, associated with reward and punishment, and with the experience of *significance* that can attach to certain objects or properties, allowing them to play a particular role in driving our reasoning. Shitij Kapur has developed an *aberrant salience* model of schizophrenia: in individuals with psychosis an excess of dopamine results in the misattribution of salience to run-of-the-mill objects or experiences. Delusions and hallucinations can be understood as an attempt to explain away that sense of significance: "Delusions are a cognitive effort by the patient to make sense of these aberrantly salient experiences, whereas hallucinations reflect a direct experience of the aberrant salience of internal representations" (2003:13).

An adult sitting in a car outside a house, for instance, would not be particularly salient to a neurotypical individual. But an excess of dopamine might result in an uncanny sense of significance being attributed to the experience of seeing such an individual. That felt need for an explanation then drives the individual to generate delusional beliefs, for instance, paranoid beliefs that they are being watched by a government agency, which explain the sense that there is something uncanny going on.

The aberrant salience account of psychosis helps demonstrate how salience can play an important role in directing our cognitive energies, shaping our patterns of reasoning and the resultant belief-set in the process. This is just as true in neurotypical individuals as for those living with schizophrenia. Salience plays an essential role in determining what abductive inferences we draw on the basis of a given body of information.

Abductive inference is mysterious in part because it ought to be impossible. In formulating an inference to the *best* explanation, the subject must consider *all other* possible explanations, which requires in turn a survey of all the available evidence to select that which is relevant to the explanation of the phenomenon in question. But creatures with finite minds cannot perform such a comprehensive survey.

Max Coltheart et al. describes this problem as it arises for Sherlock Holmes: "Holmes didn't laboriously search through every single item in memory: To be effective, his abductive inference required the use of only a small subset of his beliefs. The paradox of this global form of abduction, then, is that on the one hand, in principle "the whole background of epistemic commitments [should] be somehow brought to bear in planning and belief fixation", whilst on the other hand it is only feasible in real time to access only the relevant information" (Fodor 2001, 37). Yet somehow, as if by magic, we ordinary reasoners, like Sherlock Holmes, successfully manage this selective access, judging relevance well enough to get by, most of the time" (Coltheart, Menzies, and Sutton 2010, 276).

Where does salience feature in this process? In order for a creature with a finite mind to make these selections, information must be *ordered* by relevance to the task at hand and salience plays a

part in determining that ordering, allowing the subject to consider only a subset of the most relevant possibilities. Salient information calls out to be the focus of abductive inference. The selection of competing hypotheses will also be determined by which alternatives are salient, as will the selection of evidence against which they are assessed.[5] Successful abductive inference involves the selection of information at every stage of the process. Salience underwrites that selection process.

In order to perform these various functions, the assignation of salience must be a competitive process. Information is salient *relative to* other information. The promotion of some information comes at the expense of the demotion of other information. This aspect of salience will become important to my argument later on.

3. Prejudice and the organization of information

I turn now to three types of cases in which an individual manifests a kind of prejudice. I will argue that the most natural way of capturing what makes these instances of prejudice appeals to the organisation of information.

The core set of cases I am interested in involve what I call *perseveration*, that is, mental activity *within an existing belief set,* which need result in neither additions to nor subtractions from it. Consider the case of Margaret.

Perseveration

Margaret holds a range of true and rationally formed beliefs about Muslims, some of which are positively valenced, and some of which are negatively valenced. These include beliefs about Muslim people in her community, as well as beliefs formed on the basis of news reports, documentary films and history books. For instance, Margaret holds various beliefs her Muslim neighbours and co-workers, and also has some knowledge of the history of Islamic architecture and scholarship, and more recently of extremist Islamic terrorism and the implication of Muslim men in a case involving the sexual abuse of teenage girls in Rotherham in the UK. Margaret almost never calls to mind the positively-valenced beliefs, though she is able to access them when prompted. She frequently calls to mind the negatively-valenced beliefs.

A related set of cases concern the role of salience in determining what information a subject directs their attention towards.

Selection

Mark is idly scrolling through PhilPapers looking for literature for a paper he's working on. He scrolls until he sees a paper that looks interesting, and then reads the associated abstract to see if it's relevant. All of the papers that catch his attention are by male authors. Mark forms various beliefs about these papers. He scrolls past but does not attend to some papers by female authors. Mark does not form beliefs about the

[5] Again, clinical cases provide examples of cases in which unusual patterns of salience give rise to strange conclusions via abductive reasoning. In Capgras symdrome, subjects are convinced that a loved one has been replaced by a look-a-like imposter. Coltheart attributes the disorder to flaws in abductive reasoning:

"In delusional subjects…the balance between the whole background belief system and the particular evidence base is not successfully achieved. Capgras patients, for example, exhibit nonstandard reasoning in the sense that they do not efficiently use the right subset of background beliefs, or check their hypothesis effectively against other information available to them. As a result, their abductive reasoning isn't reliable." (Coltheart 2010)

Relatedly, recent work on illness anxiety (hypochondriasis) implicates attentional biases in the generation of health concerns (Owens et al. 2004), and hypothesizes that unusual patterns of attention to afferent information from within the body may drive inferences that support the anxiety (Ainley et al. 2016).

papers with female authors.

In this case, certain papers jump out at Mark as meriting further investigation. These are all authored by men. This systematic patterning in the distribution of his cognitive resources gives rise to a situation in which he systematically neglects papers by female authors. That reflects the competitive nature of salience: the salience of papers written by men precludes his attention to papers written by women, without him having to actively seek to avoid that information, or even process it at all. When it comes to the selection of information, salience plays a role in determining what new beliefs an individual acquires.

The third set of cases involve abductive inference. As we saw above, abductive inference relies on the selection or prioritisation of information, so it is unsurprising that salience plays a role in shaping abductive inference in ways that can lead to prejudicial outcomes.

> *Abduction: Educational Psychologist*
> Amy, a white educational psychologist, is assessing an African American child's academic performance. Amy performs a battery of tests whose results are complex and ultimately ambiguous. The child's race is highly salient to Amy, and in turn makes salient to her both the possibility that the child will be less cognitively able, and certain test results on which the child performed less well. Amy is more confident as a result than she would otherwise be that the child has certain cognitive deficits. In fact, the child does have these deficits. Had Amy attended instead to certain other test results she would have formed the contrary conclusion, which would have been false.

Salience can also be an important driver of abductive inference that leads to false conclusions. Consider the case of Frank:

> *Abduction: Disagreement at Work*
> Frank is an older man who has worked for many years in the same job. His company hires a younger woman who openly identifies as gay to work alongside him. In a departmental meeting, she publicly disagrees with a decision Frank has made. Frank experiences the disagreement as unduly aggressive, and wonders why his colleague behaved in this way. The woman's sexuality is particularly salient to Frank. He reflects on the fact that gay women have rarely been friendly to him in the past. He begins to develop a theory that gay women tend to be less friendly and more aggressive, and uses this theory to explain his encounter at work.

Part of what is significant about the preceding two cases is the way in which the salience of certain demographic characteristics means that they present themselves as *explanatory handles* for abductive inference, as the *kind* of category that might give leverage on the questions that Amy and Frank are faced with: why did this woman behave in this way, and what is this child's educational potential?

4. First pass: undue attention

In all these cases, the believer is prejudiced in some way. In some of these cases, those involving abductive inference, they may arrive at false or irrational beliefs. But that need not be the case. The problem in the *Selection* case lies as much in what beliefs are *not* formed, as in any flaw with the beliefs that *are* formed. In the case of *Perseveration,* there need be *no* change in the individual's belief

set at all, in the sense of new beliefs added or old beliefs lost. And when the subject does arrive at false or irrational beliefs, those features are explained by another, more basic feature of these cases: the way in which the subject *organises* information at their disposal.

This may require us to abstract from any realistic scenario. It is *highly likely* that all of the individuals in these cases would also harbour explicit problematic beliefs, negative emotive attitudes and behave in problematic ways towards members of the relevant group. My point is that *even if we abstract away* from that, we find that they *still* seem to be prejudiced. I propose that we treat that as an important signpost to a more general fact about prejudice.

What is it that makes these instances of prejudice? Something they all have in common is that they involve the ordering of information, and that the prejudice the subjects manifest resides in that ordering, not in the *content* of their attitudes per se. Is there anything distinctively morally or epistemically problematic about these forms of prejudice?

Since the subject's mindset in these cases is not always distinguished by obvious inaccuracy or irrational processes of belief formation, the resources of traditional epistemology are limited in their ability to pinpoint the problem here. They tend to target transitions *among* beliefs, and it is harder for them to evaluate the *failure* to form a belief, or the activity of thinking per se, that is, the tokening of a thought, when that doesn't expand or refine the contents of the subjects' belief set. *Ex hypothesi* Margaret and Mark's beliefs are broadly accurate, so evaluative tools rooted in accuracy or reliability will struggle to locate a problem here. Is Mark failing to proportion his beliefs to the evidence? Not obviously: rather, he is gathering evidence in the *wrong way*. The evidentialist, primarily oriented towards considering whether a subject's beliefs are appropriately sensitive to the evidence, is limited in their ability to evaluate practices of *gathering* evidence (Connee and Feldman 2004; Derose 2000)

More promising for this purpose is a growing body of recent work that focuses on the question of when a subject can reasonably stop looking for evidence, and the ways in which a subject may be culpable for prematurely stopping in their search for evidence, or for failing to acquire evidence they could have had, but failed to have (see, for instance Ballentyne (2015), Cloos (2015), Goldberg (2016), Harman (1980), Kornblith (1983) and Siegel (2013)). Though these accounts provide helpful resources to deal with the case of Mark in particular, they have less to offer us when we try to describe what makes Margaret prejudiced.[6] There is something intuitively irrational about her perseveration on a limited subset of information whose valence doesn't reflect the valence of her belief set as a whole. But how are we to describe that irrationality? What is the rational significance of *thought* when it results in no change to the belief set on which it is based?

One natural way of describing the relevant feature in these three cases is in terms of attention. All these cases involve subjects paying undue attention to the demographic properties of race, religion or gender. Mark attends only to papers by male authors. Margaret attends disproportionately to negatively-valenced information about Muslims. And Amy forms an accurate belief but only the basis of inappropriate attention to a certain subset of evidence and hypotheses. One lesson we could take from these cases is that undue attention of this kind is sufficient for prejudice. An individual is prejudiced when they are disposed to attend in this way. A belief is prejudiced when it is the result of a pattern of attention of this kind, regardless of whether it is accurate, and whether the

[6] There is however a legitimate concern that Margaret's perseveration would be likely to shape her acquisition of further evidence and inferential practice further down the line.

inferences involved in its formation are valid inferences. We could offer the following, then, as a first attempt at capturing what these cases reveal about prejudice:

Undue attention: Prejudicial attitudes can be constituted by inappropriate attention to demographic features.

But the set of cases of inappropriate attention is a highly heterogenous class. Most obviously, attention can be inappropriate because it is disproportionately directed at *negative* characteristics of a particular group, as is the case with Margaret, and her focus on negatively-valenced facts about Muslims, and her neglect of positively-valenced facts.

But the attention in question need not be to negative features in order to be undue. *Excessive* attention to demographic characteristics, even of a positive nature, can constitute a kind of prejudice. Consider the case of a Black employee, whose white colleague lavishes her afro with positive attention. Or a woman whose pregnancy prompts frequent positive comments from her colleagues, to the exclusion of her professional achievements. This kind of problematic positive attention also lies behind a certain form of microaggression: we can explain why it can be problematic to compliment someone of a different race for their competency with the English language, even when the comment is well intentioned, partly in terms of the misdirection of attention towards an expected shortcoming on the basis of that racial category.

Importantly, a *lack* of attention can be just as problematic as an excess of either positive or negative attention. It is that lack of attention that gives rise to phenomena such as white ignorance, the failure of white people to value or attend to or use information about other races, leading to a systematic pattern of ignorance about those groups and their experiences (Mills 2007). This is part of the problem with Mark: he is not attending positively *or* negatively to the work of women in philosophy. It is his *failure* to attend that is the problem.

The heterogeneity of the kinds of attention which may be inappropriate should lead us to dig deeper, to look for a lower common denominator which explains why these disparate forms of attention can all give rise to prejudice: what is it that unifies the set of patterns of attention that seem to be problematic?

Susanna Siegel argues that certain patterns of attention are problematic when they *inherit* problematic attitudes, and that we can that appraise them by the same norms that we apply to those attitudes (Siegel 2017, Ch.9). On her view, patterns of attention can be driven by inferences: inferences regulate what information a subject takes in – for instance, when a subject infers that she has collected *enough*, or *all the relevant* information. Inferences transmit warrant, but they also transmit epistemic flaws. If a problematic attitude is what drives an inference, then that problem is passed down to the output of the inference – in this case, the pattern of attention. As a result, "patterns of attention are appraisable, either by moral or epistemic norms, when they inherit an outlook that is itself appraisable by those norms." (2017, 160).

This move offers us powerful resources for evaluating patterns of attention. It raises two further questions for our purposes. One is whether information gathering is governed by an inference of the relevant kind in all relevant cases. If it is not, then we have reason to continue to look for other features that unify the set of cases in which attention is problematically distributed. The second question, or challenge, is to identify the problematic attitude driving the relevant inference, and then

to give an account of *what that consists in.* The underlying challenge – of understanding the nature of that prejudicial attitude – is pushed back. On Siegel's account, that problematic attitude is distinct from the resulting pattern of attention. I want to suggest that rather than *inheriting* an outlook, patterns of ordering information – attentional and otherwise - can themselves be *constitutive* of such an outlook. Given the heterogeneity of the various patterns of attention that can exemplify a form of prejudice, we need to dig deeper, to look at the salience structures which in turn drive these patterns of attention. It is to these that I turn in the next section.

5. Unpacking Salience Structures

Diverse patterns of attention can give rise to prejudice. What, then, is the common factor underlying prejudicial patterns of attention? The answer I offer here is that they are all driven by salience structures that are organised to an undue degree around demographic categories. Indeed, I claim that such salience structures can be *constitutive* of prejudice. But I first need to explain what I mean by a salience structure.

If we think of an individual as in possession of a set of beliefs, then we can understand salience as the way in which that set is organised, so as to render some beliefs more accessible than others.[7] Just as the way in which a cupboard is organised renders some objects in it easier to access than others, so salience determines the ease with which a thinker can draw on or move between pieces of information. In Margaret's case, negative beliefs about Muslims are more accessible than more positive beliefs. As a result, these beliefs serve as magnets for her abductive processes. Just knowing the *content* of someone's belief set, without knowing how that set is organised, leaves out a lot of what matters about their mental landscape – facts which will determine how they reason about and relate to the world.

This is a good start, but it will not quite do for our purposes. A structuring that scopes just over our existing beliefs will only include information that is already *within the ken* of the individual in some form. Part of what our examples show, however, is that salience doesn't only structure information already within the mind, but information that is yet to be acquired. Salience shapes not just the space of answers to a question that the reasoner has in mind, but also the space of possible questions that they are able or likely to ask. To capture that we need to think of salience as shaping not just a belief set, but rather a possible space of information, that includes both information already accessed by a subject, and information they have yet to acquire.

Salient information is more accessible, both when it is information that a subject already has, and when it is information that they have yet to acquire.[8] Information about the red postbox is more accessible than information about other features of the street corner, even before you have looked at the area in question. Information about a topic that is rendered salient by positive or negative affect is similarly easier to access when reasoning or conversing. Ease of access in these cases can be understood in terms of the cognitive effort the subject would have to expend in order to access the

[7] In the account that follows, I focus on the salience *of information,* and set aside therefore the very significant role of attention in generating *action.* Wayne Wu (2011) offers a rich account of attention as operating primarily in a *behavioural* space, where it is required to solve the "Many-Many Problem" of generating coherent behaviour on the basis of many inputs and many possible outputs.

[8] My account of a salience structure closely equates salience with accessibility. See Higgins (1996) for a discussion of the historical association in psychology of the notions of salience and accessibility and a nuanced consideration of the way in which the two can come apart.

information in question, where information is 'accessed' when it is actively drawn upon by a cognitive or perceptual system or subsystem. In this sense, then, when Mark opens the PhilPapers website, it would take a greater cognitive effort for him to access the information about papers written by female authors, because that information is less salient to him. Taken together, these facts about accessibility determine his salience structure.

A salience structure is thus in part determined by thoroughly mundane facts about the organisation of the subject's environment. If there are fewer books written about Black film directors, or if those books are only available in specialist bookshops, then that information is less accessible to the subject that information about white film directors. We therefore need to consider not just the cognitive effort that a subject has to put in to accessing that information, but also the run of the mill *physical* effort of conducting an enquiry in a given environment. But it is also often helpful to distinguish the role these two factors play in determining accessibility: even if some books are readily to hand, the information in them may remain less accessible as a result of features of the individual's interests or their social context.[9]

In placing salience structures at the centre of an account of this form of prejudice, we should be reassured by the way in which the notion of a salience structure, or close relatives of it, have been fruitfully drawn on across a range of disciplines. Sebastian Watzl develops the closely related idea of a *priority structure* to describe the ways in which salience determines how we assign our attention to our environment (Watzl 2017). Cognitive linguistics appeals to salience as a way of capturing the attention-attracting properties of some entities and concepts, and the ways in which world formation and acquisition is shaped by well-entrenched mental structures (Schmid 2010). Cognitive psychology posits topographically-arranged saliency maps that represent the visual salience of a scene (Koch and Ullman 1987). Those maps have then been used in artificial intelligence to help generate synthetic vision (Potapova, Zillich, and Vincze 2017), and to facilitate object identification on the basis of eye-tracking (Papadopoulos et al. 2014). Eye-tracking reveals what information is valued by or salient to the subject. From that, artificial intelligence is capable of identifying what the item of interest is.

But there are also some significant points of diversion between these notions of salience, and the salience structures I appeal to. Firstly, my salience structures are an abstraction, realised by a combination of psychological processes, operating over a range of psychological objects, and facts about the social and physical environment. Those work in conjunction to determine how accessible information is to a given individual.

Secondly, and relatedly, my concept of a salience structure scopes primarily over *information* rather than *psychological states*. Salience structures are multiply realisable: the information that they scope over can take many forms. Often it will be encoded in a psychological state. These include propositional states like belief, supposition or imagination, but also conceptual associations and

[9] We could think of a salience structure as akin to a map of the internet, in which links among pages are represented as lines connecting up different nodes. This resultant map shows how some webpages are very well connected: they are highly accessible, in the sense that you are likely to visit them in the course of any enquiry, even if expending little or no deliberate effort to do so. These websites are like highly salient pieces of information. Less well-connected websites, which receive little traffic and which you would have to search for deliberately in order to access, are like less salient pieces of information. How accessible a website is comes apart from the quality of information the website contains. Some websites might be highly accurate but very inaccessible, and vice versa. So it is in an individual: how well-formed a belief is, or how accurate an element in a visual experience is come apart from how accessible those states are.

perceptual experiences. Information that the subject has yet to access will, of course, not yet be encoded in such a state, but it can still play a part in determining their salience structure. An individual's mind may move among these pieces of information via a range of mental processes, including association and inference. Information is non-factive: false beliefs encode information, it just fails to correspond to how the world actually is.

Arising from this, the notion of salience structures that I appeal to is thoroughgoingly *externalist,* unlike the salience structures drawn on in cognitive psychology and linguistics. Accessibility relations are determined in part by facts about the mind of the individual, but they also depend on facts about the individual's social and physical environment.

Finally, on my account the structure itself is distinct from the propositional content of the beliefs themselves.[10] That is important for the following reason: a salience structure orders information in a way that renders certain abductive inferential tasks computationally tractable.[11] They allow that a subject need not consider every possible hypothesis or belief: instead they can just take the top few pieces of information from a pile that is pre-ordered. If the ordering were simply contained *in* the propositional content of the state in question, the subject would still need to become aware of the content of each possible belief before they could know which ones were the most relevant for a given tasks. This would not solve the problem at hand.[12]

The accessibility of information shifts depending on task: context and goal render certain pieces of information more accessible than others. But underneath those shifting sands we can identify a default ordering that evolves much more slowly over time, determined in part by a chronic calcification on the basis of repeated task performance, stable contextual features, and standing individual interests and goals. For instance, if I am trying to decide what I should buy my brother for his birthday, certain information will be made more or less accessible by the particular constraints of that task and the conditions under which I am performing it (it is raining and it is late, but a nearby bookshop is still open), but standing interests and feelings towards my brother will also play a role in determining the accessibility ordering that holds at that moment (his long-term interest in rabbits, for instance, makes rabbit-related information and present ideas particularly salient).[13] It would be an immensely complex task to distinguish the underlying ordering from the shallow influence of these local, contextual features, particularly since, over time, the local organisation

[10] This is a significant point of departure from what Watzl's "priority structures". In his book, *Structuring the Mind* , Watzl argues that attention is regulated by priority structures, where salience consists in a certain kind of imperatival content – an instruction to put something at the top of the priority structure (Watzl 2017). On my account salience cannot be reduced to additional content.

[11] See Stroud (1979) and Gardiner (2012) for further arguments that structural relations among beliefs cannot reduce to further beliefs.

[12] In this sense my model is more closely aligned with Elisabeth Camp's notion of frames. Frames, according to Camp, have an important non-propositional aspect. You don't have to be able to spell out the features of their weighting in order to have the relevant frame (Camp 2020).

[13] This echoes the notion of entrenchment, common in models of cognitive salience in linguistics. According to Hans-Jörg Schmid "Cognitive units come to be entrenched and their activation automated to the extent that they have been used before" (Schmid 2010, 118). Ronald Langacker describes a "continuous scale of entrenchment in cognitive organization. Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence." (Langacker 1987, 59).

impacts on the underlying structure.[14]

In summary, then, a salience structure can be understood as a default accessibility ordering over a space of possible information. A salience structure solves a frame problem: information comes pre-ordered by accessibility. Salience structures can be driven by the projects and interests of the individual, but importantly they are also acquired passively, and are significantly determined by an individual's social and physical environment.

6. Salience and Prejudice

With this model of salience structures in hand, I am now in a position to provide the following definition of a *prejudicial* salience structure:

A prejudicial attitude towards a demographic group can be constituted by a salience structure which is unduly organised around that demographic category.

What is it for a salience structure to be organised around a category? And when is such organisation *undue*?

Organisation around a feature

Return to the metaphor of a landscape. We could say that a landscape is organised around a particular feature when that feature heavily determines other facts about the landscape: routes of access between points on it, for instance, or the placement of certain features. The area around Snowdon is organised around that mountain: a map with no contour lines will still reveal the prominence of the mountain because the arrangement of other features of the map and routes of access between them are determined by its presence. We have to appeal to that organising object in order to explain other features of the landscape.

A salience structure is arranged around a category when that category determines the accessibility relations between information in the salience structure. That can be revealed by the role of that category in explanations of the relevant accessibility relations. This may become apparent when we appeal to the category to explain why it was particularly hard for someone to access a given piece of information, or why a certain visual stimulus was particularly salient: "Of course I noticed the sign right away because it had a photograph of a schnauzer on it!" More generally, a salience structure is organised around a particular category if we have to appeal to that category to explain and predict the patterning of their accessibility relations. Whilst those organising categories will change depending on the task in question, underlying those changes is a default accessibility ordering which reflects longer term categorical priorities.

[14] As Camp puts it in her description of frames: "Altering the prominence or centrality of a single feature can often induce pervasive complex alterations to the structural relations among other elements by 'tipping' them into new clusters of explanatory and other dependence relations and new weightings of prominence" (Camp 2020, 312).

When is such organisation undue?

The claim that the organisation of information around a category can be undue implies that the organisation of information is normatively assessible. I will not here provide a full account of all the ways in which this assessment can be performed but will just say enough to support the key claim that we can give an account of a certain kind of prejudice in these terms. Information can be more or less *rewarding* for the subject. Information can be rewarding in different ways. It is *practically* rewarding when it helps them to achieve their practical goals, such as finding food or seeing themselves as a reasonably likable person. It is *epistemically* rewarding when it reduces uncertainty. On another axis, information is *idiosyncratically* rewarding when it allows the subject to realise practical or epistemic goals that are idiosyncratic to the individual, and it is *objectively* rewarding when those goals are ones that are independent of any individual's particular idiosyncratic preferences or projects.

Undue arrangement around a given category *manifests* or *is detectable* when that arrangement serves to hinder the subject from achieving certain goals, be they idiosyncratic goals, or objective goals attributed to them by a 3rd party, such as rational thought or action. In identifying prejudice we need to appeal to objective rather than idiosyncratic goals. Within a structurally racist or sexist society, for instance, prejudicial salience structures can actually facilitate the individual in achieving their idiosyncratic goals or reducing their subjective uncertainty.

A prejudicial salience structure may hinder the pursuit of a number of objective goals: practical, epistemic and ethical. These will provide overlapping and sometimes conflicting standards of evaluation, and disagreement about whether something counts as prejudice or not can sometimes be understood in terms of disagreement about the relevant or most important goals relative to which that evaluation should be made.

We do not here need to decide which of these goals is the main or crucial one relative to which prejudice is identified, but I want to identify one particularly important goal: the epistemic goal of acquiring information, (where that can be quantified in terms of the resolution of maximal uncertainty). This is a background goal that can be hampered by many things: the need to stop reading and eat a sandwich, for instance. But prejudicial salience structures obstruct us from this goal in a particular way: through their organisation around a demographic category. Conceptual priming, for instance, is prejudicial when it has the result that information which would more efficiently resolve uncertainty is overlooked in favour of information that is already known. Many features of a salience structure can have this effect - an obsession with a football team, for instance, that leads to a particular organisation of information. What makes this a case of prejudice is that it is organisation around a *demographic* category that has the relevant effect.

It might be the case that prejudicial salience structures sometimes result in highly efficient acquisition of information (when you quickly notice that a woman is slender, for instance), but this will generally be only in some very specific domain. The specificity of the benefit is outweighed by the epistemic costs in other respects: your failure to notice other things about the woman in question.

7. Prejudice Lite?

What I have offered is a very *minimal* account of prejudice. Prejudice doesn't require a particular animus against a group, (except in so far as the arrangement of information can itself be regarded as constituting such an animus). There is no requirement of dislike, or intent to harm. An individual

need not be capable of consciously accessing or endorsing a particular attitude, nor need we attribute implicit *belief* states to them. In this section I want to explore two potential objections to that minimality, before drawing out some of the resources it offers us.

This may seem to some to be too *lite* an understanding of prejudice. Won't this account predict - overpredict indeed - a lot of internalised prejudice? Since prejudice can be constituted by a salience structure, and since salience structures can be acquired passively, through exposure to culturally common associations, for instance, then many members of marginalised groups will have a salience structure that reflects those societal prejudices. Moreover, this structure of associations may coexist with explicit antiprejudicial attitudes.

To a limited extent, this is a bullet that we should bite. Being a woman, even a committed feminist, is no protection against misogynistic salience structures. If it is harder for someone to access information about a woman's opinions or skill set than it is information about the distribution of fat on her body, even when the latter is irrelevant to the task at hand, then there is prejudice there no matter how right-on her explicit belief set is. This is a significant way in which internalised prejudice persists: it is easier to reject explicit attitudes, or to consciously wish good things for others than it is to alter the default ways in which we order and access information. In recognising that as a form of prejudice we need not claim that it is amongst the most egregious but it is important nonetheless.

This might seem like a particularly problematic verdict when we consider that absorbing a dominant salience structure may in fact be a prerequisite for a certain kind of anti-prejudicial awareness. Consider a Black anti-racist activist. Being able to see through the lens of a racist salience structure is crucial to their successful activism: they have to understand and appreciate the ways in which the dominant ideology operates, and the information it makes available. Similarly, a committed feminist will be alive to the ubiquity of sexism, but to do that she will inevitably adopt aspects of the dominant, prejudicial salience structure. How are these people to operate effectively *without* absorbing a prejudicial salience structure?

At this juncture I want to object that these individuals simply *do not have* the same salience structure as someone with a straightforwardly prejudicial attitude. Rather, their awareness of the oppressive conditions under which they operate lets them access quite a different set of information than would the prototypical racist or sexist. They explicitly notice certain accessibility relations, but that awareness simultaneously changes their salience structure. Recognising their own tendency to order information in line with that dominant, prejudicial salience structure itself makes certain other pieces of information accessible: in recognising how readily they register information about the size and shape of a woman's body, information about the operation of patriarchal power structures is also made accessible, in ways it is not to someone with a prejudicial attitude. In this way, their salience structure has *some* features in common with a prejudicial salience structure, but not all. Its organisation around those demographic categories is not undue in the same straightforward way that a prejudicial salience structure is. It may be organised around those categories but in a way that is more proportionate, and that ultimately helps rather than hinders the acquisition of information.

A prejudicial salience structure is one that is unduly organised around a demographic category. In the previous section I recommended that we understand undue organisation relative to an epistemic standard, but other more practical standards also present themselves, and it is worth noting here that a salience structure that is epistemically costly may nonetheless be practically

advantageous, even a necessity for survival. Recognising this lets us appreciate the dilemma that structurally prejudiced societies present their members with: there is no way of maximizing pragmatic and epistemic goods. Practical survival can come at an epistemic cost, and those practical stakes are highest for members of oppressed groups: salience structures are one way in which that operates.

A second important objection complains that this account puts the cart before the horse. The problem in these cases isn't the salience structure itself, it's the *attitude* which drives it. The reason we feel that the salience structure in these cases is problematic is that it's driven *by prejudicial animus*. Mark doesn't attend to information about female authors, because on some level, however hard to access it might be, he *thinks less of them* or cares less about them. Margaret perseverates on negative information about Muslims *because she doesn't like them*. That's where the problem lies, not in the relative accessibility of information.

This inclination to identify some more basic attitude which itself constitutes the problematic prejudicial state is natural, but it is to be resisted. Of course, we can identify other, richer forms of prejudice which require more by way of substantive mental states. But the power of the account offered here is that it avoids the requirement that we *reify* prejudice as a *state* or *feeling,* a positive, identifiable attitude of the individual that we can point to. It is this reification which mandates unproductive investigations into what someone's attitude towards a particular group *really* is, in cases where behaviour and belief, or explicit and implicit attitudes collide. On the approach I advocate, we do not need to divine what Amy's emotions towards African Americans *really* are, nor weigh them against her beliefs and behaviour to arrive at some final reckoning. The arrangement of information constitutes a kind of prejudice whether or not it is accompanied by some more substantive racial animus.

In some respects this amounts to a rejection of the categories offered by more standard approaches. But the account also has the power to cash out some of the terms or constructs they appeal to. Take an alternative construal of Margaret's case, that claims that she is prejudiced because her thought about Muslims is organised around a *stereotype,* and reliance on stereotypes is constitutive of a kind of prejudice. It is natural to then ask what *is* a stereotype? And what is wrong with using them? One way of understanding what a stereotype is, is *just as* a particular way of organising information, a rigid set of accessibility relations, such that recognizing that someone is a Muslim makes information about terrorism or female repression highly salient, or less harmfully, that makes information about tea and the royal family more accessible when someone speaks with a British accent. What is wrong with stereotypes is that they make other, inconsistent information much harder to access. The apparent explanatory power of appeal to stereotypes should not rely just on the obscurity of that notion. Cashing stereotypes out in terms of a salience structure accommodates our urge to appeal to them in identifying certain forms of prejudice whilst also offering us a deeper explanatory handle on what is wrong with them

Finally, it is worth nothing that this account is well placed to offer an account of when ignorance is prejudicial and when it is not, by allowing for a distinction between *active* ignorance, and *passive* ignorance, where the first, like white ignorance, may be constitutive of prejudicial attitudes in a way in which the latter is not. Active ignorance arises when a salience structure is organised around a particular category, so as to render information about that category systematically less accessible. Passive ignorance is when information about a category is inaccessible not as a result of the

organisation of the relevant structure, but just as a question of happenstance. My ignorance of facts about mitochondria is of this latter sort. I don't have that information, but not as a result of the systematic organisation of my salience structure around that category. Appealing to the category of mitochondria does not help to explain or predict features of my cognition more broadly. By contrast, white ignorance arises when information is rendered inaccessible because of the organisation of the individual's salience structure around racial categories: those categories are highly predictive of the relevant accessibility relations. As Charles Mills puts it, White ignorance is "a pervasively deforming outlook - that was not contingent but causally linked to their whiteness" (Mills 2015, 217).

8. Conclusion

The organization of information is an essential feature of the mind, and plays a central role in determining not just the inferences a subject draws, but what evidence they encounter as well as acquire, and the kinds of thought processes they engage in, whether or not those expand their belief set, or merely hold it stable. I have appealed to the notion of a salience structure - a set of accessibility relations that encompass information that is already known to the subject and information they have yet to encounter, to allow us to capture the way in which this prioritisation of information can constitute a form of prejudice. If correct, this account has a wide range of more practical implications. We should explicitly consider what factors shape an individual's salience structure, and wat we can do to influence them in anti-prejudicial ways. This will have implications for the design of on-line environments, for instance: what information a webpage prioritises and what it deprioritises matters, as well as the explicit content that it carries. Part of media literacy involves noticing how media outlets seek to manipulate one's salience structure in prejudicial ways, literally highlighting the cellulite on a woman's thigh, for instance, or by eye-catching visual images that conform to certain stereotypes.

Works Cited

Ainley, Vivien, Matthew A. J. Apps, Aikaterini Fotopoulou, and Manos Tsakiris. 2016. "'Bodily Precision': A Predictive Coding Account of Individual Differences in Interoceptive Accuracy." *Philosophical Transactions of the Royal Society B: Biological Sciences* 371 (1708): 20160003. https://doi.org/10.1098/rstb.2016.0003.

Appiah, Kwame Anthony. 1990. "Racisms." In *Anatomy of Racism*, edited by David Goldberg, 3–17.

Ballantyne, Nathan. 2015. "THE SIGNIFICANCE OF UNPOSSESSED EVIDENCE." *The Philosophical Quarterly* 65 (260): 315–35. https://doi.org/10.1093/pq/pqu096.

Blum, Lawrence. 2004. "What Do Accounts of 'Racism' Do?" In *Racism in Mind*, by MIchael Levine and Tamas Pataki. Ithaca: Cornell University Press.

Camp, Elisabeth. 2020. "Imaginative Frames for Scientific Inquiry: Metaphors, Telling Facts, and Just-So StoriesMetaphors, Telling Facts, and Just-So Stories." In , 304–36. https://doi.org/10.1093/oso/9780190212308.003.0014.

Cloos, Christopher Michael. 2015. "Responsibilist Evidentialism." *Philosophical Studies* 172 (11): 2999–3016.

Coltheart, Max, Peter Menzies, and John Sutton. 2010. "Abductive Inference and Delusional Belief." *Cognitive Neuropsychiatry* 15 (1–3): 261–87. https://doi.org/10.1080/13546800903439120.

Connee, Earl, and Richard Feldman. 2004. *Evidentialism: Essays in Epistemology*. Oxford University Press.

Corlett, J. Angelo. 1998. "Analyzing Racism." *Public Affairs Quarterly* 12 (1): 23–50.

Cudd, Ann E., and Leslie E. Jones. 2007. "Sexism." In *A Companion to Applied Ethics*, 102–17. John Wiley &

Sons, Ltd. https://doi.org/10.1002/9780470996621.ch8.
Derose, Keith. 2000. "Ought We to Follow Our Evidence?" *Philosophy and Phenomenological Research* 60 (3): 697–706. https://doi.org/10.2307/2653824.
Fodor, Jerry A. 2001. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. MIT press.
Garcia, J. L. A. 1996. "The Heart of Racism." *Journal of Social Philosophy* 27 (1): 5–46. https://doi.org/10.1111/j.1467-9833.1996.tb00225.x.
Gardiner, Georgi. 2012. "Understanding, Integration, and Epistemic Value" *Acta Analytica* 27 (2): 163-181
Gaspelin, Nicholas, and Steven Luck. 2018. "The Role of Inhibition in Avoiding Distraction by Salient Stimuli." *Trends in Cognitive Sciences* 22 (November): 79–92. https://doi.org/10.1016/j.tics.2017.11.001.
Glasgow, Joshua. 2009. "Racism as Disrespect." *Ethics* 120 (1): 64–93. https://doi.org/10.1086/648588.
Goldberg, Sanford C. 2016. "On the Epistemic Significance of Evidence You Should Have Had." *Episteme* 13 (4): 449–470.
Greenwald, Anthony G, Debbie E McGhee, and Jordan L K Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology*. 74 (6): 1464–1480.
Harman, Gilbert. 1980. "Reasoning and Evidence One Does Not Possess1." *Midwest Studies in Philosophy* 5 (1): 163–182.
Higgins, Edward Tory. 1996. "Knowledge Activation: Accessibility, Applicability, and Salience." In *Social Psychology: Handbook of Basic Principles*, 133–68. New York, NY, US: The Guilford Press.
Humphreys, Glyn W., and Jie Sui. 2015. "The Salient Self: Social Saliency Effects Based on Self-Bias." *Journal of Cognitive Psychology* 27 (2): 129–40. https://doi.org/10.1080/20445911.2014.996156.
Johnson, Gabbrielle M. forthcoming. "The Structure of Bias." *Mind*. https://doi.org/10.1093/mind/fzaa011.
Kamke, Marc R., and Jill Harris. 2014. "Contingent Capture of Involuntary Visual Attention Interferes with Detection of Auditory Stimuli." *Frontiers in Psychology* 5. https://doi.org/10.3389/fpsyg.2014.00528.
Koch, Christof, and Shimon Ullman. 1987. "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry." In *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience*, edited by Lucia M. Vaina, 115–41. Synthese Library. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-3833-5_5.
Kornblith, Hilary. 1983. "Justified Belief and Epistemically Responsible Action." *Philosophical Review* 92 (1): 33–48.
Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. Stanford University Press.
Levy, Neil. 2017. "Am I a Racist? Implicit Bias and the Ascription of Racism." *The Philosophical Quarterly* 67 (268): 534–51. https://doi.org/10.1093/pq/pqw070.
Lien, Mei-Ching, Eric Ruthruff, Zachary Goodin, and Roger W. Remington. 2008. "Contingent Attentional Capture by Top-down Control Settings: Converging Evidence from Event-Related Potentials." *Journal of Experimental Psychology. Human Perception and Performance* 34 (3): 509–30. https://doi.org/10.1037/0096-1523.34.3.509.
Mandelbaum, Eric. 2016. "Attitude, Inference, Association: On the Propositional Structure of Implicit Bias." *Noûs* 50 (3): 629–58. https://doi.org/10.1111/nous.12089.
Mills, Charles. 2007. "White Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan Nancy Tuana, 11–38. State Univ of New York Pr.
———. 2015. "Global White Ignorance." Edited by Matthias Gross and Lindsay McGoey. *Routledge International Handbook of Ignorance Studies*, 217–227.
Moody-Adams, Michele. 2007. "Racism." In *A Companion to Applied Ethics*, 89–101. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470996621.ch7.
Owens, Katherine M. B., Gordon J. G. Asmundson, Thomas Hadjistavropoulos, and Travis J. Owens. 2004. "Attentional Bias Toward Illness Threat in Individuals with Elevated Health Anxiety." *Cognitive Therapy and Research* 28 (1): 57–66. https://doi.org/10.1023/B:COTR.0000016930.85884.29.
Papadopoulos, Dim P., Alasdair D. F. Clarke, Frank Keller, and Vittorio Ferrari. 2014. "Training Object Class Detectors from Eye Tracking Data." In *Computer Vision – ECCV 2014*, edited by David Fleet, Tomas

Pajdla, Bernt Schiele, and Tinne Tuytelaars, 361–76. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_24.

Potapova, Ekaterina, Michael Zillich, and Markus Vincze. 2017. "Survey of Recent Advances in 3D Visual Attention for Robotics." *The International Journal of Robotics Research* 36 (11): 1159–76. https://doi.org/10.1177/0278364917726587.

Ramsey, Frank. 1990. "General Propositions and Causality." In *F. P. Ramsey: Philosophical Papers.*, edited by D. H. Mellor. Cambridge University Press.

Schmid, Hans-Jörg. 2010. "Entrenchment, Salience, and Basic Levels." *The Oxford Handbook of Cognitive Linguistics*, June. https://doi.org/10.1093/oxfordhb/9780199738632.013.0005.

Siegel, S. 2013. "Can Selection Effects on Experience Influence Its Rational Role?" In *Oxford Studies in Epistemology*, edited by T. Gendler, 4:240. Oxford.

Siegel, Susanna. 2017. *The Rationality of Perception*.

Stroud, Barry. 1979. "Inference, Belief and Understanding" *Mind* 86:179-196

Watzl, Sebastian. 2017. *Structuring Mind. The Nature of Attention and How It Shapes Consciousness.* Oxford, UK: Oxford University Press.

Wu, Wayne. 2011. "Attention as Selection for Action." In *Attention: Philosophical and Psychological Essays*, edited by Christopher Mole, Declan Smithies, and Wayne Wu, 97–116. Oxford University Press.