YOU ARE NOT WORTH THE RISK: THE ETHICS OF STATISTICAL DISCRIMINATION IN

ORGANISATIONAL SELECTION OF APPLICANTS

BY

VANESSA SCHOLES

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

Victoria University of Wellington

2017

Abstract

Your job application is rejected unseen because you ticked a box admitting you smoke. The employer screened out applicants who ticked the 'smoker' box, because she had read empirical studies that suggest smokers, as a group, are a higher productivity risk than non-smokers. What distinctive ethical concerns inhere in the organisational practice of discriminating against applicants on the basis of group risk statistics? I argue that risk-focussed statistical discrimination is morally undesirable due to the lack of respect for applicants as unique autonomous agents. However, I argue further that the decision-making context affects the morality of this discrimination. Other things being equal, the morality of statistical discrimination varies depending on the purpose of the organisation, the level of detail in the discrimination, and whether the discrimination is transparent to applicants and includes some benefit for applicants. Because organisations may have good reason to use risk-focussed statistical discrimination when assessing applicants, I present some recommendations for decision-makers to mitigate the lack of respect for applicants as individual agents. Organisational decision-makers can focus on the extent to which the statistical data they use comprise i) factors that feature efforts and achievements of the applicant; ii) dynamic rather than static factors; and iii) data drawn from the applicant's own history and actions over time.

Acknowledgements

I would like to acknowledge first and foremost the support of my primary supervisor, Dr Ramon Das. Ramon's willingness to listen, his collegial encouragement and helpful advice and support for my goals enabled me to start and complete my thesis and some publications over the last four years. I would also like to acknowledge the support of my secondary supervisor, Professor Simon Keller. Simon's timely feedback was very valuable in helping me improve my work. Thanks to James Cone, my office partner Caroline McDonald and Andrew McCaw for help with reading through chapters or discussing thesis progress – much appreciated. Finally, I owe a big debt of gratitude to my parents, Lorraine and Geoff Scholes, who supported me with care (and childcare) during the course of my study.

Table of Contents

CHAPTER FIVE **Using learning analytics to risk-screen applicants to tertiary education**

5.1 Risk-focussed learning analytics

5.2 Conceptions of the ethics of learning analytics

5.3 Discrimination and learning analytics

5.4 Group risk statistics and individual students

5.5 Recommendations

CHAPTER SIX   **Statistical discrimination in hiring**

6.1 Discussing statistical discrimination in employment

6.2 Types of statistical discrimination in hiring

6.3 Comparison with traditional hiring methods

6.4 The information the process seeks

6.5 What the job requires

6.6 Recommendations

CHAPTER SEVEN   **Concluding comments**

7.1 Legislation to address statistical discrimination

7.2 Conclusion

7.3 Summation

Chapter one   Introduction

Suppose that an employer knows that some neighbourhoods have a higher risk of producing worse performing employees than other neighbourhoods. The employer has a vacancy for a job, for which workers need to have a minimum level of certification. The employer advertises the vacancy and receives a large number of responses. After removing the applicants who do not have the requisite minimum level of certification, he is still left with a large number of applications. He sorts the job applicants on whether they live in a 'good' or 'bad' neighbourhood, and rejects all the applicants from 'bad' neighbourhoods. Now he has a more manageable pile of applications to consider.

This thesis will examine the ethics of organisational decisions to select or reject applicants for relevant goods based on risk statistics. This is a form of statistical discrimination where an organisational decision-maker uses risk statistics for a group to assess individuals, and treats individuals from 'riskier' groups worse. Of particular moral significance are discriminations that serve to exclude applicants without further consideration. Suppose that the aforementioned employer had a computer program to screen applications and automatically reject applicants from the riskiest neighbourhoods. Applicants are denied any consideration for relevant goods due to a statistic for a group of which they are a member, regardless of their suitability as an individual.

There are several ways in which differences between groups can make it potentially beneficial for an organisational decision-maker to screen applicants using statistical discrimination to control for risk. When I refer to statistical discrimination I will have one of the following three grounds in mind. One basis for discrimination is where there is a difference in the *average* for one group compared with another group. Suppose that Tamara runs a property management service that sources tenants for rental houses. She knows there will be a large pool of applicants, as there is a housing shortage. Tamara has read research suggesting undergraduate students are worse at keeping to a regular rent payment schedule than postgraduate students. In her online application form, she includes a couple of tick boxes that pop up if an applicant chooses 'student' as their occupation, that forces the student to identify as either undergraduate or postgraduate, and automatically rejects applicants that specify 'undergraduate'.

A second basis for statistical discrimination is where the average for two groups is the same but there is a higher *variance* in the members of one group compared with another (Albert & Cabrillo, 2000, p6). Imagine a city with a spread of suburbs, many of which have relatively homogenous sub-populations. Suburb X, however, has two distinct sub populations. Two thirds of Suburb X comprises well-resourced families whose young adult members are thereby *more* financially secure than populations in surrounding suburbs in the city. The remaining third of Suburb X is on the wrong side of the tracks; it comprises gang families on low and intermittent income streams, whose young adult members are thereby far *less* financially secure than the populations in surrounding suburbs. The average financial security measure for Suburb X as a whole is no different from any other suburb; but there is a wider variance in the financial security of suburb members. Suppose that Tamara requires a family guarantor for applicants for tenancy who have little credit history. *Ex hypothesi*, if Tamara selects such an applicant from Suburb X rather than another suburb she has a higher chance of getting a tenant whose guarantor is more financially secure, but also a higher chance of getting an applicant whose guarantor is significantly less financially secure. If Tamara is risk-averse, she would have an incentive to choose a member from another suburb over a member of Suburb X.

A third basis for statistical discrimination involves variability in the results for groups due to the test or selection process, rather than due to significant variability in the groups themselves (Albert & Cabrillo, 2000, p8). This is where the test—or tester—for predicting performance produces more accurate predictions for one group than for another[1]. For instance, suppose that Tamara gets a large number of applications for a few vacancies. She decides to do a quick sort of the applications looking at the quality of the credit history information provided, to whittle down the number. She is familiar with a reasonable number of the organisations with whom applicants have credit histories. But the organisations mentioned by some applicants are from other countries and are unfamiliar to

---

[1] For a detailed discussion of this issue, see Schmidt and Hunter (1976).

her. She does her quick initial sort of applications. When the credit histories featuring recognised organisations are sorted, some go in the 'Yes' pile, and some in the 'No' pile (depending on the quality of the credit history respectively). But for the applications featuring 'unfamiliar' organisations, *all* go in the 'No' pile. Here the benefit to Tamara's organisation from the discrimination is not due to any difference in average credit history of applicants with familiar versus unfamiliar organisations, but from the avoidance of the uncertainty associated with the latter. She can produce a better statistical assessment for the familiar organisations.

The thesis addresses practices and policies of statistical discrimination that are applied to persons applying to organisations for particular goods. The goods examined will include insurance, a job, and admission to tertiary education. Group features selected for scrutiny will be those associated with risk for the goods under consideration. If the relevance or impact of a feature is believed to be unclear, decision-makers are faced with an uncertainty that stymies its use. If a probability can be determined for the group feature, this creates a rational basis for statistical discrimination. Probabilities of this sort are discovered through empirical research. Assigning a probability effectively turns an uncertainty into a situation of quantifiable risk – thus creating a measurable incentive for engaging in statistical discrimination. So it is to be expected that the very pursuit of empirical research into correlations between group features and particular goods will prompt increasing use of statistical discrimination in the future (see Gandy, 2010). That is, research into which features correlate with lower job performance or tertiary education failure lays the groundwork for the use of statistical discrimination in these areas.

My key question is: what distinctive ethical concerns inhere in the practice of discriminating against applicants on the basis of group risk statistics? I argue that statistical discrimination is a morally problematic form of discrimination. I defend this claim in part by critically considering some current analyses of the practice of statistical discrimination by philosophers. In particular, I argue that the distinction between statistical evidence and individualised evidence is sharper than some analyses suggest, with implications for the autonomy and agency of applicants in organisational selection decisions. Insofar as

13

statistical evidence in these decisions automatically displaces individualised evidence, decision-makers cannot treat an applicant as an individual, which denies the applicant's unique autonomous agency. If we think a person is morally valuable (at least in part) as a locus of unique moral agency, this is morally concerning.

Being able to treat individual applicants as autonomous agents is morally important. However, there may be competing ethical considerations for organisational decision-makers that favour statistical discrimination, such as the consequences for everyone affected. Bearing this in mind, I argue that several factors influence the morality of statistical discrimination in organisational selection of applicants. I argue that the moral status of this discrimination varies according to the purpose of the organisation, which I illustrate by contrasting the purpose of courts of justice with that of insurance institutions. Another factor is whether the applicants subject to the discrimination gain any benefit. A third factor is the transparency of the discrimination, which affects whether applicants can be held to consent to it. Lastly, the moral status of statistical discrimination varies according to the extent to which the statistical data used is drawn from the subject's own history and actions over time rather being drawn simply from data on groups of other people. The last two factors affect whether interactions between the organisational decision-maker and the applicant respect the individuality and agency of persons.

In this chapter, I begin by distinguishing statistical discrimination from the more familiar 'taste'-based discrimination. In section 1.2 I outline my methodology, explaining my choice of a pluralist, non-ideal theory approach. In section 1.3 I comment on how statistical discrimination fits – or often does not fit – a selection of philosophical accounts of wrongful discrimination, mostly from the last decade. I then turn to an article by Oscar Horta (2010) on the concept of discrimination *simpliciter*. In section 1.4 I argue, contra Horta, that disadvantageous exclusion on the basis of group generalisations is central to discrimination. I also argue that our concept of discrimination requires that it be *a potentially wrongful exercise of a legitimate practice* – a criterion met by the statistical discrimination I will discuss. Section 1.5 concludes this chapter with a brief overview of the chapters to come.

1.1 How statistical discrimination differs from 'taste' discrimination

Philosophers' accounts of wrongful discrimination have tended to focus on 'taste'-based discrimination, which involves treating people differently on the basis of a personal taste or prejudice against a particular group. Alexander (1992, pp158-167) distinguishes two broad forms of 'taste'-based discrimination, based on categorical or contingent preferences. A categorical preference is a thorough-going bias regarding a group; for example, someone who is biased against an ethnic group would try to avoid the group across all contexts in their life, such as work, study, friendships, and home life. Bias is thus a wholesale aversion to a group, reflected in things like misogyny and homophobia. A contingent discriminatory preference, by contrast, is a preference to avoid a group in a particular context only, not across all contexts. For example, a person might have no aversion to working or studying with people of a particular ethnicity, but would be averse to having a person of that ethnicity as a partner or intimate friend. Alternatively, a person might have female friends, but be prejudiced against having a female boss. Or, someone might be fine interacting with males in almost all contexts, but averse to the idea of men focusing on a primary childcare role.

Statistical discrimination differs significantly from taste-based discriminations, by the decision-maker's motive and their rationale (their conscious justification for the action). For example, suppose that a manager has read studies indicating that curiosity is correlated with job performance[2]. The manager decides to include an online psychometric test for curiosity in her hiring process and reject applicants categorised as below average in curiosity. The manager's motive for discriminating does not involve any personal distaste for the less curious or support for a general social prejudice against less curious applicants. Instead, the manager's motive is to increase the productivity of the workforce, and thereby increase profits for the firm. Similarly, the manager's *rationale* is not one of ignoring a group of people with a characteristic she personally finds less appealing, nor that this group should be restricted in their employment opportunities. Instead, her rationale could be to carry out a fiduciary responsibility to shareholders to increase profits through legally permissible

---

[2] See e.g. Mussel, P. (2012) Introducing the construct curiosity for predicting job performance. *Journal of Organizational Behavior*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/job.1809/full

actions; and it is not illegal to discriminate against the less curious. So the sorts of motives and rationales that we might find ethically objectionable for taste-discrimination do not (or do not necessarily) apply in the case of statistical discrimination.

Kasper Lippert-Rasmussen offers the following formal definition of statistical discrimination:

> A policy, P, constitutes statistical discrimination against a certain socially salient group of people, X-people, in relation to non-X-people (or some subgroup thereof) if, and only if, (i) there is statistical evidence which suggests that X-people differ from non-X-people in dimension D, (ii) P involves treating X-people worse than non-X-people, and (iii) P is in place because of (i). (Lippert-Rasmussen, 2007, p387)

Lippert-Rasmussen restricts his focus in this definition to a "socially salient group" which is treated worse though statistical discrimination. Lippert-Rasmussen suggests a group is socially salient if it is recognised as a group by society, and the grouping structures social interactions in a variety of contexts. This covers broad groupings such as by gender, ethnicity, nationality and religion. Whether this covers a group such as 'science fiction fans' is less clear. In any case, in my next chapter I will argue against restricting our consideration to a "socially salient group", but otherwise I think Lippert-Rasmussen's definition of statistical discrimination is good.

The practice of discriminating between people on the basis of statistical generalisations about group features without personal prejudice has been referred to as 'discrimination by proxy' (see Alexander and Cole, 1997), because the decision-maker can be using a group feature as a proxy for information about a person. For example, knowing that smokers, as a group, suffer worse health outcomes than non-smokers, an employer might use smoking as a proxy for 'poorer health and greater productivity risk'. The practice has also been referred to as 'rational discrimination' (see, e.g., Harford, 2008), because it is not based in (irrational) prejudice and because it serves the prudential interests of the decision-maker. However, the practice is most commonly referred to as 'statistical discrimination', as this is how it is

known in the economic literature, which is where it has had most discussion[3]. This is also the term employed by Lippert-Rasmussen, the philosopher who has given the most attention to the practice recently in the philosophical literature. Accordingly, 'statistical discrimination' is the term I use in this thesis.

1.2 Methodology

It is probably not common for doctoral theses in philosophy to include a methodology section, so-called. Indeed, it seems philosophers are not prone to discussing their methodologies (McBain, 2012, p1). I think applied ethics is a special case, because there are (at least)[4] two broad categories of methodological issues commonly addressed in the field. The ideal vs. non-ideal theory distinction addresses what sort of question an applied ethicist hopes to answer, and determines methodology from there. The top-down vs. bottom-up distinction addresses the preferred approach to analyzing the moral issues inherent in the question. I will discuss these distinctions briefly in this section, and identify where my thesis stands in regard to these. In term of my overall methodological approach, I aspire to a form of pragmatism identified by Koopman (2012, p5)[5] as characterised by "fallibilism, pluralism, naturalism, antiskepticism, and antifoundationalism, as well as by a practice of philosophy that is at once interdisciplinary and committed to its public relevance".

Introduced by Rawls (1971), the categories 'ideal theory' and 'non-ideal theory' respectively distinguish between a theory that aims to set out what the best principles would be were everyone to follow them, and a theory that aims to take account of the state of the real world, where it is not the case that we can expect everyone to follow these principles. There has been increased interest in this distinction over the last few years (Levy, 2014). What makes a theory non-ideal is the allowance that the just thing to do might not embody the ideal conditions or follow the rules set out in an ideal theory. This is because non-ideal

---

[3] The term 'statistical discrimination' is also used in sociology; for example, see Baumle and Fossett (2005).
[4] Chadwick and Schroeder (2002, p2) suggest that "Methodologies are as varied as the questions applied ethicists face. Some, such as reflective equilibrium, virtue ethics or discourse ethics, have been developed in traditional areas of philosophy. Others, such as principlism (Volume II) or casuistry have emerged in particular areas of applied ethics."
[5] Koopman is discussing Kwame Anthony Appiah, suggesting his approach to applied ethics fits these pragmatist criteria.

theory does not assume that the real world is such that actions or policies expressing the principles of an ideal theory will necessarily produce justice in the real world. Hamlin and Stemplowska (2012, pp48-49) note that in contrast to theorising perfect or ideal justice (or goodness, or some other value), non-ideal theory can aim for local improvement in justice (or goodness, or some other value). Stemplowska and Swift (2014) argue that if the road to implementing the conditions for an ideal theory is very long or hard to navigate, these can be reasons to focus at the level of non-ideal theory. This is because trying to get to the end of a very long road may allow too much injustice or suffering in current generations, and because we might have greater expectations for effectiveness in aiming for local improvements.

My thesis involves non-ideal theorising because of the assumptions I make about the organisational decision-makers who are choosing whether or how to practice statistical discrimination in the selection of applicants. I assume that these decision-makers are individuals, or organisationally-bound groups making a decision as an individual (such as a management group), operating in a real-world context over whose conditions they have little influence. For example, if there are existing laws or competitive pressures, the decision-maker has to take these into account, and it will not be reasonable to expect her to be able to change these. I therefore eschew solutions that are only actionable for legislators or well-funded lobby groups. The geographical context for my organisational decision-makers should be read as Western, English-speaking developed countries, such as the United States, Canada, the United Kingdom, Australia and New Zealand. I hope my work is also relevant to non-English speaking developed countries; however, my research primarily draws on books and articles from researchers from English-speaking developed countries, and the aforementioned countries in particular.

My interest is the identification of the ethical concerns proper to the real-life social practice of risk-based statistical discrimination in selection of applicants to organisations. In order to identify ethical concerns, there are three reasonably distinct methodologies that could be employed, and choosing one could mean quite a different project than if another were chosen. The first methodology is to take a 'top-down' approach of applying a normative ethical theory to assess a particular real-life issue (Beauchamp & Childress, 2008; Dittmer, 2013). For example, Stephen Maitzen (1991) takes a 'top-down' consequentialist approach

to statistical discrimination. He applies a principle of utility to derive high-level principles for statistical discrimination in social policies such as race-based affirmative action programs or sex-based pension plans (p23). The second methodology is a 'pluralist' approach (Dittmer, 2013) or 'coherence theory' approach (Beauchamp and Childress, 2008) of applying mid-level moral principles or values to assess a real-life issue. In this approach, a researcher can use a principle, such as "do no harm", or a value such as autonomy or equality of opportunity. The third methodology is a 'bottom-up' approach of starting from the full particulars of a specific real-life situation presented as a case study (Beauchamp & Childress, 2008; Dittmer, 2013). With the 'bottom-up' approach, also referred to as casuistry (Chadwick & Schroeder, 2002, p2), researchers must decide whether to compare cases, and if so, what cases to compare.

There are benefits and drawbacks for each applied ethics methodology. Taking a 'top-down' approach risks alienating the significant proportion of the philosophical audience that rejects the chosen normative ethical theory. This proportion is likely to be around two-thirds of philosophers, according to Bourget and Chalmers' (2009) survey.[6] The same general concern applies for a 'coherence' approach that specifies in advance the general moral principles or values to be used. However, this will be mitigated to the extent that the principles or values are relevant (for different reasons of course) in more than one major normative ethical theory. For example, "do no harm" is relevant to utilitarians as in some situations doing no harm may be the way to minimise suffering; it is also relevant for virtue ethicists as doing no harm may be required to avoid cruelty. For the 'bottom-up' approach, there is a risk of producing analysis so particular to the case at hand that it does not generalise at all (Beauchamp & Childress, 2008). My preference is to produce a thesis whose analyses are potentially useful for and relevant to a wide philosophical audience, so my chosen approach is the pluralist 'coherence' approach to examining statistical discrimination, informed by investigations of statistical discrimination in relation to the real-life practices of hiring and admission to tertiary education.

---

[6] Bourget and Chalmers' 2008-2009 PhilPapers survey of professional philosophers' views shows each of three major ethical traditions (consequentialism, deontology, virtue ethics) achieving less than one-third assent from participants (Bourget & Chalmers, 2013).

Applied ethicists may assess the concepts people commonly use to evaluate real-life moral action, such as, for example, injustice, corruption, discrimination or equality. Applied ethicists may also directly assess real laws, practices, policies, institutions (being a special set of practices) and particular actions. Kopelman (1990, p21) suggests that we ought to conceive of the latter assessments as changing, or having the potential to change, any substantive principles or background theories employed in the endeavor of applied ethics. The genuine possibility exists that "[w]hatever we employ to solve problems might be changed, bolstered, specified, clarified, altered, rethought, or challenged when used" (p201). I agree with Kopelman's view. Accordingly, I start with a provisional definition of a social practice as ethically concerning, beyond its immediate overall consequences in terms of happiness or harms, if it reduces people's opportunities to choose to engage in social practices as the particular individuals they are, or are becoming. I will draw on general moral values such as welfare and autonomy to assess the ethics of the practice of statistical discrimination, and be open to refining or redefining these principles in light of consideration of the practice. This means the application of the principles will be determined partly through what seems salient when analysing the practice in real life.

A further question regarding methods for a project in applied ethics is the extent to which a practitioner identifies and draws on relevant empirical research. This is part of a broader epistemological question regarding the nature of the 'particulars' that will be used as a base for analysis. Will evaluation from particulars proceed primarily from introspection about how things are 'at the bottom' (perhaps informed by anecdotes or engagement with other philosophers)? Or will reasoning proceed from relevant existing empirical research on the particulars? Christina Bicchieri (2006) argues that philosophy needs to take into account empirical research from social sciences such as psychology, as what people ought to do is bounded by what they are able to do or what it is reasonable to assume they can do. Bicchieri (2006, p23) concludes that "…if philosophy wants to thrive as a normative enterprise it will have to integrate those scientific results that give it a better grounding." I will draw on relevant empirical evidence from secondary sources where I can to inform my analysis. This can be seen, for example, in chapter two where my discussion of stigma and stereotypes will be informed by research from psychology, sociology and economics.

Finally, I agree with Kaiser, Millar, Thorstensen and Tomkins (2007, p68) that ethically sound frameworks for investigating practical concerns ought to encompass five features. First is the inclusion of values at stake. In the next part of this chapter I note other philosophers' moral concerns with discrimination and how these may be included in assessing statistical discrimination, and in chapters 2 and 3 my discussion supports individuality, autonomy and agency as particular values at stake. The second feature is transparency: making clear the background thinking and reasons on which views are based. I hope this methodology section aids transparency; in the other chapters I support transparency through using examples and sometimes visuals to explain ideas, and clarifying when claims are supported by empirical research. The third feature is a multiplicity of viewpoints. I address this through including other philosophers' views; viewpoints from psychology and economics, and both the applicants' perspective and the organisational agents' perspective on organisational selection of applicants. The fourth feature is the exposition of case-relevant ethically-relevant aspects; this is covered in chapters 5 and 6, and includes identification of existing practices and relevant laws. The last feature is to include ethical arguments, which I do throughout the thesis.

1.3 Philosophical accounts of wrongful discrimination

There is no philosophical consensus on how to specify the moral wrong(s) of discrimination. As noted earlier, many widely cited contributions from philosophers focus primarily on defining the wrong of 'taste'-based discrimination. Larry Alexander (1992), for example, sees this form of discrimination as failing to correctly acknowledge a person's moral equality or moral worth in a way that harms others; and Richard Arneson (2006, p779) argues wrongful discrimination is defective conduct based on "unjustified hostile attitudes toward people perceived to be of a certain kind or faulty beliefs about the characteristics of people of that type". These accounts of discrimination, while distinct, cluster around the concept of differential treatment of individuals based in morally offensive attitudes or beliefs about group characteristics. Statistical discrimination, by contrast, has nothing to do with attitudes or beliefs about the moral worth of groups. Because it is distinct from taste-based discrimination, it is perhaps not surprising that statistical discrimination does not connect

with these definitions. However, other accounts of discrimination can encompass statistical discrimination, and I will look at these for help in framing moral questions on the topic.

Deborah Hellman's (2008) account of discrimination as a form of demeaning a person suggests the moral wrong of discrimination lies in what the discriminatory act expresses. She claims that neither the intention of the agent nor the consequences of the act are necessary to defining what the discriminatory act expresses. Instead, the expressive nature of the discriminatory act is a function of the social context giving its denigratory meaning coupled with the agent being in a position of power relative to the person subject to the discrimination. 'Demeaning' acts or policies express a disregard for the moral equality of those being discriminated against. Hellman's definition covers 'taste' and 'indirect' discrimination; the moral wrongness of a particular instance of either form of discrimination depends on whether we appropriately understand that instance to be demeaning. Her definition will presumably count statistical discrimination morally wrong if it somehow brings into question the moral equality of those subject to it. This may seem unlikely; as noted earlier, this discrimination does not appear to pass judgement on moral worth. Nevertheless, I will still ask the general question: what does statistical discrimination to deselect applicants seem to express about people, and is this morally problematic?

The next two accounts of discrimination I will mention are directly consequentialist. Kasper Lippert-Rasmussen (2006) gives an account of the wrongness of private discrimination, focussing on discrimination by individuals rather than by a state. Lippert-Rasmussen's ethical view has its basis in consequentialism (pp818-9); he thinks the badness of any discriminatory act is due to its bad consequences, or due to its being one of a set of such acts that together produces bad consequences. However, he qualifies this view (2006, pp833-4) with the inclusion of prioritarian, desert-accommodating rules. His basic suggestion is that acts are morally good the more they increase well-being; where there's a choice, we should prioritise the well-being of the worse-off; and where there's a choice, we should prioritise well-being for those that deserve it. I will consider the consequences for welfare, and distribution of these consequences, in my analyses of statistical discrimination.

Shlomi Segall (2012) argues that what is bad about discrimination is its consequences for our equality of opportunity with regard to our welfare. Discrimination is bad when it

undermines this equality. Segall says (2012, p83) his concept of equality of opportunity is much broader than, for example, one person lacking the same opportunity as someone else applying for a job. Instead, it is about the level of equality across the various opportunities available in society to improve our welfare, of which a job is only one example. On this view, it is not obvious that a practice of discrimination in one area, such as hiring, can be assessed in isolation from other practices impacting on equality of welfare opportunity. These could include, for example, mortgage lending or insurance, but also access to social welfare, access to tertiary education, and access to health care, among others. An assessment of statistical discrimination at this level would be a project in political science, and beyond the scope of my thesis. However, I take from Segall's view the importance of noting the extent to which discrimination that limits equality of welfare opportunity in one area is compounded by - or mitigated by – discrimination in other areas.

The last two theorists I will mention present moral concerns with discrimination that are closest to my worries about statistical discrimination. Sophia Moreau (2010) presents an indirectly consequentialist approach to what makes discrimination wrong, defining it as differential treatment that has negative consequences for our 'deliberative freedoms'. 'Deliberative freedoms' are freedoms to access core life opportunities without the concern that one's 'normatively extraneous' traits will be counted as costs against oneself. 'Normatively extraneous' traits are those that we think people should not be evaluated on in the distribution of opportunities to access some goods and services - traits such as gender or skin colour or religious tradition. Moreau's definition would cover statistical discrimination on the basis of such traits. It might also cover statistical discrimination for characteristics such as height and residential demographics, for example, whilst probably leaving out characteristics such as smoking – dependent on which characteristics we think fit her idea of 'normatively extraneous'. Even if some characteristics are not covered, I will consider statistical discrimination's impact (if any) on people's subjective autonomy and liberty to access social goods.

Moreau (2013, p136) writes that "… who you are is, at least in part, the choices you have made; and when you face discrimination, these choices are unfairly constrained both by other people's assumptions about extraneous traits of yours and by their failure to accommodate these traits." This is echoed somewhat by Benjamin Eidelson (2013). His

23

concern (p395) is that discrimination fails to treat a person properly as an individual, "... by failing to treat him as in part a product of his own past efforts at self-creation, and as an autonomous agent whose future choices are his own to make." The difference between the two theorists lies in how each would unpack Moreau's term 'unfairly constrained' in relation to autonomy. Moreau would focus more on the *consequences* of the discrimination, the fact that the person does not get to make their choices freely. Eidelson would seem to focus more on the discriminatory *practice* itself, the fact that a person is not recognised as an individual in the practice.

The claim that discrimination is morally problematic because it fails to treat someone as an individual person arises in ordinary discourse on discrimination and in the legal rhetoric of courts (Eidelson, 2013, p355). However the claim has not received much attention in the literature, aside from the work of Lippert-Rasmussen (2007, 2011) who is fairly critical of it. Eidelson, however, argues that we should see it as part of the badness of discrimination due to the importance of individuality to our autonomy. This sounds somewhat consequentialist, but Eidelson seems to suggest rather that the morally problematic aspect is attitudinal or epistemic: it is about not properly acknowledging someone's dignity as a person, not according them proper moral recognition. Eidelson's concern would suggest many instances of statistical discrimination are in some sense bad, as statistical discrimination would seem to be a paradigm case of failing to treat people as individuals. This will be a key focus in my thesis. I will argue that where disadvantaging or excluding people on the basis of group risk statistics is morally wrong, the wrong (at least in part) lies in a failure to treat people as individuals.

1.4 On the necessity of a group criterion to our idea of wrongful discrimination

As noted, statistical discrimination lacks some elements we might see as salient for making discrimination morally bad, such as bearing ill-will toward the person against whom one discriminates, or expressing animosity toward a group of persons. If we omit any mention of groups from accounts of discrimination with those elements, there still seems to be behaviour of moral concern; namely, bearing ill-will or animosity toward persons. But

statistical discrimination features the imposition of a group criterion on applicants without (or at least, without requiring) any such biased behaviours. A group criterion is a feature that makes someone a member of some group. Anything that is of distinctive moral concern with the practice of statistical discrimination will not be about hostility toward people, but will instead be centrally bound up with assessment that imposes a group criterion on individuals.

Philosophers discussing wrongful discrimination have tended to simply proceed on the basis that a group criterion is part of our concept of discrimination (see Thomsen, 2013, p121). This acceptance has recently been challenged by Oscar Horta (2008). Horta argues against retaining any group criterion for the concept of (wrongful) discrimination, instead proposing a definition of discrimination as treatment of a person that is (basically) disadvantageous and unjustified. If this were an appropriate definition of discrimination, then rational statistical discrimination might not count as a form of discrimination at all: while disadvantageous to a person, it is not (or at least not necessarily) unjustified. I will argue that there are flaws with Horta's position and that we should retain the use of a group criterion as a basic part of our idea of discrimination[7].

Horta argues for a formal definition of discrimination that can be used to discern when an individual is inadequately considered in the application of a normative moral criterion. "It often happens that different individuals are not treated equally with regard to what a certain normative criterion prescribes. With some qualifications, …. it would be very useful to use the concept of "discrimination" in these cases." (Horta, 2010, p314). He raises the point (2010, p314) that despite a great many attempts to identify when discrimination occurs, there are relatively few definitions of "discrimination" in the literature. He notes the following lines of definition: a mere preference for someone (Block & Walker, 1982); differential consideration (Boxill, 1991); disadvantageous enjoyment of rights (Sawer, 2000); inequality of opportunity (Nagel, 1991); and the idea that the disadvantageous consideration or treatment must be related to group membership in order to be discriminatory (Wasserman, 1998; Lippert-Rasmussen 2006, 2007; Cavanagh, 2002; Young

---

[7] Note that my argument is only against Horta's position on the necessity of a group-criterion for discrimination, rather than his arguments in other parts of his article.

1990). Horta suggests that many disagreements over when discrimination occurs are simply disagreements over the meaning of 'discrimination'.

Horta states (p316) that all the aforementioned definitions agree with the broad claim that discrimination takes place when some individuals are treated or considered in a way that is disadvantageous for them. He notes this claim does not cover three aspects we want a definition of (morally relevant) discrimination to cover: 1) intention (say where an agent makes a credible attempt to discriminate that would normally result in disadvantage but is foiled by something that means the 'victim' does not end up with any disadvantage); 2) those cases where we think the disadvantageous treatment is justified; 3) where the agent is forced / seriously coerced to do an action that disadvantages someone but had no discriminatory intent (pp316-18). In light of this, Horta argues for a more complex but still broad definition of discrimination as taking place when "some individuals are treated or considered in a way that is disadvantageous for them" (p316) "[and] unjustified" (p317); "[and] for some disadvantageous treatment to be discriminatory it needs to be unjustified according to first order reasons" (p320).

Horta recognises that his account does not connect the unjustified disadvantageous treatment with membership in a socially salient group (p316). He notes that some theorists would disagree with this; they would say that disadvantageous consideration of a person as an individual (not based in the individual's group membership) is not discrimination. Horta suggests that perhaps these theorists' disagreement with his view is "merely linguistic" based on the theorists' cultural context (p316). He writes:

> But how are we going to decide which [disadvantageous unjustified acts] should
> be called "discrimination"? This dispute cannot be decided by an appeal to
> natural language, because the term is differently understood to this respect by
> different people. Other arguments may be introduced here, then. (Horta, 2010,
> p317)

However, Horta has not sufficiently established that discrimination *is* understood differently by different people, when it comes to the group criterion. It is not just Wasserman, Lippert-Rasmussen, Cavanagh and Young who hold a group criterion to be part of discrimination; Boxill, Nagel and Sawer all conceive of discrimination as a phenomenon applying to

members of particular social groups[8]. Moreover, Edmonds (2006) argues that our idea of discrimination is fundamentally underpinned by the concept of social identity-groups or 'castes' (identity-groups).

Block and Walker (1982) certainly present a different understanding, but their sweeping definition of discrimination as the simple expression of a preference for one thing over another does not capture discrimination in a moral sense. It is true that discrimination in its broadest form can refer to a 'telling apart' of some thing from another thing. Suppose that someone shows you two blue cards that are extremely similar in hue and says to you: "Many people see these as exactly the same colour. Can you discriminate between them?" If you truthfully answer yes, you have discriminated. However, at this level, your discrimination is not morally assessable activity. Discrimination as a morally assessable activity requires making a choice in relation to moral subjects, such as people, where your choice is significantly influenced by the positive or negative value you attached to the subject/s.

Being charitable, we could focus on Block and Walker's definition just in regard to preferences concerning people (as Horta does when he cites Block and Walker). However, this definition would still include a vast number of considerations or treatments of persons that do not seem to connect to the moral sphere (for example, "I prefer Mozart to Beethoven" or "I prefer to tie my own shoelaces rather than hire someone to do this"). As Block and Walker (1982, p.1) note "[s]trictly speaking … the term is neutral in application". But Horta does not want to appeal to a sense of 'discrimination' that is so neutral in application that it does not pick out ethically-relevant discrimination, even if he does want to remain neutral between particular normative ethical theories.

That said, Horta could still claim that it is better for some reason to define discrimination more broadly - which he does thus:

---

[8] Boxill's article is about affirmative action with regards to racism and sexism. Nagel's chapter does not actually define discrimination, instead talking of the inequality of opportunity that is caused by it; in any case, his writing clearly indicates he sees discrimination as involving disadvantage based on social grouping. Emeritus Professor Sawer's (2000) definition of discrimination is written for the purpose of describing and explaining the particular discrimination that is a global issue faced by women.

> … the more general definition I have presented here has at least one advantage. Although it may render the term "discrimination" less useful for those dealing with the disadvantageous consideration of some social groups, it makes it possible for us to use it as a general concept to examine whether some criterion for moral consideration is being rightly applied. (Horta, 2010, p317)

Horta (2010, p317) claims his broader definition of discrimination means it can be used for the purpose of conceptualising whether, in relation to some normative theory, someone is treated or considered disadvantageously in an unjustified way. This is the main reason Horta gives for wanting to defend the broader definition and hence the thesis that "if two individuals are treated in exactly the same disadvantageous and unjustified way, the fact that only one of them belongs to a certain salient group cannot be a reason to say that she is the only one who is discriminated against"(p317). Horta notes that having to add in factual considerations about individuals' group membership would negate the usefulness of 'discrimination' for this purpose. We would then need to make up another concept to cover what happens when, in relation to some normative theory, someone faces disadvantageous, unjustified exclusion. However, Horta does not ask if we have such a concept already. On the face of it, we do; namely, 'unfairness'. To see why 'unfairness' is preferable to 'discrimination', consider the following comparison.

At the start of Horta's article, he explores how his preliminary definition of discrimination fits examples of actions we might intuitively consider discriminatory; but he does not notice that his definition covers many types of morally-relevant actions which we ordinarily (intuitively) do not count as discriminatory. Examples of such types of actions include promise-breaking, lying, bribery, disloyalty, slander, theft, assault, betrayal and murder. Horta could try to claim that these actions do discriminate against the victim, but this seems counterintuitive. A serial killer certainly treats some individuals in a way that is disadvantageous to them and unjustified, but it is not intuitive to say that the serial killer discriminates against her victims. A robber who preys on elderly folk simply because they are usually easier to steal from, certainly subjects his victims to an unjustified disadvantageous treatment, but it seems wrong to say the robber discriminates against the elderly.

However, 'unfairness' seems able, or at least better able than 'discrimination', to apply to the types of unjustified disadvantageous actions just mentioned, such as promise-breaking, bribery, theft and murder. While we will surely want to add other terms to describe the moral wrongness, at least it does not seem counterintuitive to say in each of these cases the perpetrator has acted unfairly toward the victims[9]. Horta owes us some explanation of why 'unfairness' is less satisfactory a term than 'discrimination'. Until then, the position taken by most theorists that disadvantageous treatment needs to be related to group membership in order to be discriminatory seems preferable.

It may be fruitful to compare the earlier unjustified disadvantageous actions with those we intuitively think of as discriminatory. If we can describe a general difference that distinguishes the latter, we will have made progress on specifying the nature of moral discrimination. Think of some obvious cases of discriminatory action, such as racist or sexist hiring practices. Now think of the disadvantageous cases of theft and murder. A feature distinguishing the former cases from the latter is not immediately apparent. Perhaps the latter group might usually have worse consequences, but this will not always be so. Clearly murder is almost as bad as it gets in terms of consequences; but a small-scale theft might have fewer bad consequences than, say, a Black person having to search for twice as many months to find employment.

To better focus our attention, we can compare racist or sexist hiring practices and racist or sexist murder. What distinguishes the first lot of actions from the second? I suggest that it is whether the base practice is legitimate. To speak plainly, hiring is a legitimate practice; murder is not. While refusing to let accommodation to Muslims or to the elderly is not legitimate, the underlying practice of renting out accommodation is legitimate. With, say, stealing from Muslims or making then breaking promises to the elderly, neither the base practice of stealing nor that of deceitful promising is legitimate. Hence we can tentatively propose the following as a condition of (wrongful) discrimination that circumscribes it relative to unfairness: discrimination involves *a potentially wrongful exercise of a legitimate practice*.

---

[9] Describing these types of actions as 'discriminatory', however, seems to trivialise what is of concern about those actions. At the same time, it also seems to disdain what is of concern about more typical instances of discrimination.

Horta could potentially object that actions such as promise-breaking, theft and murder do not obviously involve an agent acting to exclude someone from a benefit offered to others, which is the sort of sphere where we recognise discrimination. But I think even within this sphere, Horta's definition is weakened by not including a group criterion. For example, suppose that an employer disliked his ex-brother-in-law, Brendon. Brendon's sister applies for a job, and the employer refuses to hire her simply because he doesn't like his ex-brother-in-law and doesn't want to have anything to do with the family of his ex-brother-in-law. While certainly unfair, this seems more like a case of reverse-nepotism than a case of hiring discrimination. However, this would seem to qualify as discrimination on Horta's account.[10]

I think we can illuminate the difference by reversing the action. Suppose that an agent acted *advantageously* toward an applicant, on the same basis that they might otherwise have disadvantaged them. Applied to the example above, the employer hired the ex-brother-in-law's sister simply because he liked his ex-brother-in-law and wished him well. This would be a clear expression of nepotism. In other cases where the criterion also was not group-based but was idiosyncratic to the person being hired, the advantageous treatment would be an expression of the employer's personal preference (e.g., "I just liked her combination of positive vibe and ideas for what she could bring to the company"). But if we take cases of discrimination and reverse them, we get *bias* or *affirmative action*, rather than personal preference or the specific complaint of nepotism. If an employer hires mostly men because it is mostly men who apply for the positions, but would hire any woman over any man who did apply, that would be bias. It's not about the specific woman, but rather that the applicant fits the group criterion of being a woman.

In conclusion, I suggest we should retain the use of a group criterion as a basic part of our idea of discrimination. This is important to my thesis because, as noted earlier, anything of distinctive moral concern with the practice of statistical discrimination will be centrally bound up with imposing a group criterion on individuals. In particular, I will argue that

---

[10] Or if an employer refused to consider an applicant for a job because the applicant's full name was Susan Alison Tessa Anne Norris, and the employer didn't like the fact that her initials happen to spell SATAN, this would count as discrimination on Horta's specification.

disadvantaging or excluding an applicant purely using risk statistics for a group of which they are a member is a key moral concern with statistical discrimination.

1.5 Chapter overviews

I conclude chapter one by giving a brief overview of the chapters to come.

Chapter 2 – Stereotypes, theories of discrimination and statistical discrimination

It is notable that, excepting Kasper Lippert-Rasmussen, there is not much discussion of the ethics of statistical discrimination in prominent philosophical accounts of the ethics of discrimination from the last decade. My impression from the discussion that has occurred (see e.g., Lippert-Rasmussen, 2007, 2011, 2014; Hellman, 2008) is that the accurate practice of statistical discrimination brings with it no grounds for particular moral concern. My aim in chapter two is to argue for the moral relevance of statistical discrimination, drawing on empirical literature on stereotypes. I conclude that what is morally concerning about statistical discrimination in distinction from our traditional idea of discrimination (as represented in the positions of the philosophers covered in chapter two) is simply its failure to treat people as individuals, and attendant concerns for individual agency.

Chapter 3 – Statistical discrimination, treating applicants as individuals, and agency

In chapter three, I discuss arguments from three philosophers who have discussed statistical discrimination and treating people as individuals, and whose arguments challenge the idea that there is a moral concern with how statistical discrimination treats individuals. The philosophers discussed are Ferdinand Schoeman (1987), Frederick Schauer (2003) and Kasper Lippert-Rasmussen (2007, 2011, 2014). These three philosophers have presented arguments that challenge the distinction between using group-based evidence and individualised evidence when assessing people. However, I think the distinction still has some importance, and I will attempt to defend this view in responding to their arguments. I include some consideration of the power of organisations relative to individual applicants in choosing which statistical evidence is assessed. I conclude by noting the moral value in treating people as individuals, a value that seems to be denied when statistical discrimination is employed. I also suggest that whether this makes a difference to what the

agent ought to do will depend on other factors such as the purpose of the organisation and the resources the agent has available to them to influence the informational context.

Chapter 4 –The importance of context: purposes and resources

In this chapter I explore the importance of the decision-making context to the ethical permissibility of statistical discrimination in organisational treatment of applicants. I focus on two aspects: the purposes of the organisation, and the resources the decision-maker has available to them to influence the informational context. I argue that these two aspects are key contextual factors affecting the ethics of statistical discrimination. To argue the first point, I contrast the contexts of courts of justice and insurance. These are two contexts where I think we rightly vary our ethical assessments of statistical discrimination, due (at least in part) to the purpose of the organisation. To argue my second point on the decision-maker's resources, I discuss "all-or-nothing" decisions, a general decision-making context that Schauer has argued tends to support statistical discrimination. I will argue that whether a decision is or becomes an "all-or-nothing" affair can depend on the resources available to the decision-maker, so the question of resourcing is a contextual factor impacting on the ethics of statistical discrimination.

Having argued for the importance of context to the ethical permissibility of statistical discrimination, in chapters five and six I move into a more detailed discussion of the practice of statistical discrimination in two different contexts: admission to tertiary education, and hiring. My thesis assumes a status quo in terms of these practices. I think this is ok for a work in applied ethics from a non-ideal theory standpoint. However, I do note an issue of organisational power relating to how risk and the 'riskiness' of applicants are conceived, that I think societies should bear in mind when judging the ethics of organisations statistically discriminating on risk. One concern is if the use of statistical discrimination excludes consideration of other relevant information, particularly where its use limits the participation and agency of members of more vulnerable groups. A second concern is if its use forestalls institutional improvement (ways the decision maker could improve the organisation to facilitate opportunities for 'riskier' applicants).

Chapter 5 – Using learning analytics to categorise students on risk

A key difference between the education context and the hiring context is the purpose of the organisational decision-makers. Employers seek to identify 'risky' job applicants in order to exclude them, whereas educators seek to identify 'risky' student applicants in order to put in place education support initiatives for the students. Statistical discrimination in organisational selection of applicants for admission to tertiary education will occur through the use of analytics to risk-screen students. I explain the drivers for this use of analytics in tertiary education, and briefly discuss the literature on the ethics of analytics in education. I highlight the disrespect for the student's agency in the admission process, but acknowledge the overall benefit. I conclude with three recommendations to help mitigate this disrespect. These include the suggestion that the more individual behavioural analytics can be incorporated in risk-assessment of students, the more the risk assessments should be seen as recognising and including students as individual agents.

Chapter 6 – Statistical discrimination in hiring

In chapter six I discuss first the potential and drivers for risk-focussed statistical discrimination in the hiring context, along with a brief note on how this is treated in the economics literature. I then explain how statistical discrimination may be practiced in hiring, and compare this with 'traditional' hiring methods. The results of my analysis show that what distinguishes whether a particular hiring process counts as statistical discrimination is the relationship between *the information the process seeks,* and *what the job requires*. I look at the information the process seeks from an applicant's perspective, highlighting the applicant's interests and concerns. I show the disrespect for applicants as individuals and the constraints on applicants' agency. I then look at what the job requires from the employer's perspective, highlighting the employer's interests and concerns. I show the difference made by how we characterise an employer's purpose, and the constraints on employers operating in a competitive market. I conclude that it may be morally permissible for employers to use risk-focussed statistical discrimination in hiring, although this gives effect to a morally undesirable disrespect for applicants as autonomous individual agents. I end with some recommendations for employers to mitigate this, with particular mention of transparency.

Conclusion

My conclusion summarises the arguments presented in each chapter. I also take the opportunity to explain why I try to keep my focus at the level of the organisational agents deciding whether to implement risk-based statistical discrimination, rather than including a broader social policy perspective. This explanation includes a discussion of current research on legislation in the United States of America aimed at reducing the impact of statistical discrimination on applicants with criminal records. I suggest that social policy recommendations are best undertaken by social scientists, as social policy needs knowledge of local conditions and empirical investigation, rather than pure conceptual analysis.

Chapter One References

Albert, Rocio and Cabrillo, Francisco (2000) *Sex discrimination: Is the labour market a market for lemons?* Retrieved 15 September 2005 from http://www.ucm.es/info/iudem/cabrillo.DT00-4.pdf

Alexander, Larry (1992) What Makes Wrongful Discrimination Wrong? Biases, Preferences, Stereotypes, and Proxies. *University of Pennsylvania Law Review,* 141(1): 149-219.

Arneson, Richard J. (2006) What is wrongful discrimination? *San Diego Law Review, 43*: 775-808.

Baumle, Amanda K. and Fossett, Mark (2005) Statistical Discrimination in Employment: Its Practice, Conceptualization, and Implications for Public Policy. *The American Behavioral Scientist*. 49: 1250-1274.

Beauchamp, Tom L and Childress, James F (2008) *Principles of Biomedical Ethics*. (6th edition). New York: Oxford University Press.

Bicchieri, Christina (2006) Philosophy: what is to be done? *Topoi*, *25*(1-2): 21–23. doi:10.1007/s11245-006-0025-y

Bourget, David and Chalmers, David J. (2013) What do philosophers believe? *Philosophical Studies*, 1–36. doi:10.1007/s11098-013-0259-7

Chadwick, Ruth F. and Schroeder, Doris (2002) *Applied Ethics: Critical Concepts in Philosophy*. Taylor & Francis.

Dittmer, Joel (2013) Ethics, Applied. *Internet Encyclopedia of Philosophy*. Retrieved from http://www.iep.utm.edu/ap-ethic/

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York: Routledge.

Eidelson, Benjamin (2013) Treating people as individuals. In Deborah Hellman and Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp354-395). Oxford: Oxford University Press.

Gandy, Oscar H. (2010) Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems, *Ethics and Information Technology,* 12 (1): 29-42.

Hamlin, Alan, and Stemplowska, Zofia (2012) Theory, Ideal Theory and the Theory of Ideals. *Political Studies Review*, 10(1): 48–62. doi:10.1111/j.1478-9302.2011.00244.x

Harford, Tim (2008) The dangers of rational racism. In *The Logic of Life,* London: Abacus.

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, Mass., USA: Harvard University Press.

Horta, Oscar (2010) Discrimination in terms of moral exclusion. *Theoria*, 76: 314-322.

Kaiser, Matthias, Millar, Kate, Thorstensen, Erik and Tomkins, Sandy (2007) Developing the ethical matrix as a decision support framework: GM fish as a case study. *Journal of Agricultural and Environmental Ethics*, *20*(1): 65–80. doi:10.1007/s10806-006-9023-8

Koopman, Colin (2012) Pragmatist Resources for Experimental Philosophy: Inquiry in Place of Intuition. *Journal of Speculative Philosophy*, 26 (1): 1-24.

Kopelman, Loretta M. (1990) What is applied about 'Applied' philosophy? *Journal of Medicine and Philosophy,* 15(2): 199-218.

Levy, Jacob T. (2014) There's No Such Thing as Ideal Theory (SSRN Scholarly Paper No. ID 2420125). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2420125

Lippert-Rasmussen, Kasper (2006) Private discrimination: A prioritarian, desert-accommodating account. *San Diego Law Review, 43*: 817-856.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *The Journal of Political Philosophy*, 15: 385-403.

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15: 47-59.

Lippert-Rasmussen, Kasper (2014) *Born free and equal? A philosophical inquiry into the nature of discrimination*. New York: Oxford University Press.

Maitzen, Stephen (1991) The Ethics of Statistical Discrimination. *Social Theory and Practice*, 17: 23-45.

McBain, James (2012). Issue Introduction. *Essays in Philosophy*, 13, 1, Article 1.

Moreau, Sophia (2010) Discrimination as Negligence, *Canadian Journal of Philosophy,* 40:sup1, 123-149.

Moreau, Sophia (2013) In Defense of a Liberty-based Account of Discrimination. In Deborah Hellman and Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp.71-86). Oxford: Oxford University Press.

Mussel, Patrick (2012) Introducing the construct curiosity for predicting job performance. *Journal of Organizational Behavior*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/job.1809/full

Rawls, John (1971) *A theory of justice*. Cambridge, MA: Harvard University Press.

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Mass., USA: Harvard University Press.

Schoeman, Ferdinand (1987) Statistical vs direct evidence. *Nous, 21* (2): 179-198.

Segall, Shlomi (2012) What's so bad about discrimination? *Utilitas*, 24(01): 82–100. doi:10.1017/S0953820811000379

Stemplowska, Zofia and Swift, Adam (2013) Rawls on ideal and nonideal theory. In Jon

    Mandle and David Reidy (Ed.s) *A Companion to Rawls*. (Pp112-127) West Sussex, UK:

    Wiley Blackwell.

Thomsen, Frej Klem (2013) But some groups are more equal than others: A critical review of

    the group criterion in the concept of discrimination. *Social Theory and Practice*, 39 (1):

    120-146.

Chapter two   Stereotypes, theories of discrimination and statistical discrimination

Statistical discrimination can be implemented by organisations to exclude or otherwise disadvantage persons simply on the basis of their membership in a 'risky' group. On the face of it, this seems like the sort of discrimination that is worthy of our moral concern. The focus on the group seems to efface a person's individuality, and the 'risk' rationale paints the person as a threat to be mitigated. This statistical discrimination has generally been neglected in recent philosophical literature on discrimination, however[11]. In this chapter I give reasons why statistical discrimination is a morally important issue that needs further attention. My argument has two parts, one critical and one constructive. In the critical part I reject conceptions of discrimination which restrict it to actions disadvantaging certain socially salient groups. I also reject accounts of discrimination on which statistical discrimination cannot count as (potentially) morally wrong. In the constructive part I argue for the particular relevance of statistical discrimination to accounts of discrimination that can accommodate it. I draw on research on grouping people, stigma and stereotypes from economics, sociology and psychology to support my argument.

The basic structure of the chapter is as follows. First, I argue against the claim made by some philosophers (Edmonds, 2006; Lippert-Rasmussen, 2006) and assumed by many more (Thomsen, 2013) that discrimination must be against individuals as members of a disadvantaged caste or an existing socially salient group. Second, I identify some accounts of discrimination that rule out statistical discrimination as a potential wrongful discrimination, and give reasons to reject these accounts. Third, I note some philosophers whose accounts of discrimination can accommodate statistical discrimination, and explain why further consideration of statistical discrimination would be fruitful. There is a rational incentive for organisations to exclude or disadvantage persons on the basis of group risk characteristics in place of using individualized evidence. I conclude that this statistical discrimination is morally problematic at least in part for the simple reason that it disregards persons as individuals. This conclusion generally runs counter to other views in the philosophical literature on statistical discrimination, which I address in my next chapter.

---

[11] With the exception of the work of Kasper Lippert-Rasmussen (2007; 2011; 2014).

2.1 On the nature of the group subject to discrimination

Most of the discussion about the moral issue of discrimination, philosophical or otherwise, has centred on taste-based or indirect discrimination against people as members of specific social groups, such as different races, ethnicities, genders, disabilities, or sexual orientations, amongst others. Kasper Lippert-Rasmussen (2006, p829) thinks discrimination requires differential treatment of a person on the basis of their belonging (or being perceived to belong to) a *socially salient* group, and that this requirement applies to statistical discrimination. Recall Lippert-Rasmussen's definition of a statistically discriminatory policy:

> A policy, P, constitutes statistical discrimination against a certain socially salient group of people, X-people, in relation to non-X-people (or some subgroup thereof) if, and only if, (i) there is statistical evidence which suggests that X-people differ from non-X-people in dimension D, (ii) P involves treating X-people worse than non-X-people, and (iii) P is in place because of (i). (Lippert-Rasmussen, 2007, p387)

On Lippert-Rasmussen's view, a group is socially salient if it is recognized as a group by society, and the grouping structures social interactions in a variety of contexts. For example, the grouping "Muslim" has social salience but the grouping "people whose first names start with M" does not.

Frej Klem Thomsen (2013) notes that the perspective that there is a list of social groups against which actions are recognised as discrimination - a 'prohibited list' - is so pervasive in the philosophical literature that it is simply assumed without argument in the explanations or definitions of discrimination given by philosophers (Thomsen, 2013, p121). David Edmonds (2006) argues that not only is a group criterion necessary for an appropriate concept of discrimination, the group involved must be a socially *disadvantaged* group, a 'caste' with a history of discrimination: for example, the discrimination against Black Americans by White Americans, or discrimination against homosexual people.

Statistical discrimination against applicants to organisations can use features of socially disadvantaged castes and other recognized socially salient groups. However, it need not draw on these features. Statistical discrimination can instead be based on features

correlated with the end for which the decision-maker is selecting applicants. For example, if an employer aims to hire productive staff, and scoring highly on a psychometric test for the personality trait of 'neuroticism' is correlated with lower productivity, then this may be the feature used in discrimination[12]. We may recognize that the existence of a categorisation of 'neuroticism' requires a set or group-category of 'neurotic people', and the psychometric test creates a grouping of 'neurotic people'. There is, however, no socially salient grouping or 'caste' of 'neurotic people'. So statistical discrimination does not need to map applicants to a group that is currently (or traditionally) seen as socially stigmatized or disadvantaged.

Suppose that only the actions disadvantaging people as members of socially salient groups, or as members of 'castes', were eligible for consideration as discrimination of a sort that could cause us moral concern. While this would not show that the morality of statistical discrimination simply supervenes on traditional discrimination, it would seem to point in that direction. To support my claim that statistical discrimination is a distinctive concern in its own right, I want to suggest that a group member does not need to be from an existing socially disadvantaged or even currently socially salient group for their disadvantageous treatment to count as morally concerning discrimination. I think the treatment can count as discrimination as long as it is based on some kind of grouping (and so is not idiosyncratic to the individual affected[13]).

Consider the following finding from an economics paper by Sue Mialon and Seung Han Yoo (2014), 'Employers' Preference for Discrimination'. Mialon and Yoo (2014) model two groups of job applicants who must compete for limited positions. The only initial difference between workers in the two groups is the simple fact of a publicly available group identity that can be costlessly observed – otherwise the workers are identical. Let us suppose that one group has pointy heads and the other has square heads. Each worker must decide whether or not to upskill, in competition with other workers. Mialon and Yoo find that if one group of workers is more populous, employers have an incentive to discriminate against the minority group and hire applicants from the majority group. This will encourage majority group workers to invest in upskilling because there is a surer pay-off from it, and discourage

---

[12] Note this score does not mean the applicant is clinically disturbed or mentally ill in any sense, so differential treatment of the applicant will not be discrimination on the basis of disability.
[13] This distinguishes my position from that of Oscar Horta, whose argument against the use of any group criterion I objected to in chapter 1.

minority group members. This means there the number of more skilled applicants for the job increases (as the majority group is more populous), and when employers pre-select from this group, these (majority group) workers will then compete against each other to best signal their skills. The discrimination thus produces a better signalling effect from the majority group applicants allowing employers to more easily discern better applicants in this group, reducing the risk of hiring less productive applicants.[14]

Suppose that there are more pointy-head people than square-head people. Recall that initially, except for head shape and overall numbers, square-head people are identical to pointy-head people. Employers would be able to lower their overall risks by discriminating against square-head people and hiring pointy-head people. In this situation, square-head people are not a 'caste' - there is no history of discrimination. There is no pre-existing socially disadvantaged group. Yet I think it is intuitive to say that employers are discriminating against square-head workers, in the same sense of 'discrimination' we use for disadvantageous actions against members of socially disadvantaged groups[15].

Thomsen (2013, p143) rejects the view that there is any conceptual basis for restricting discrimination to members of socially salient groups. He suggests other aspects of discrimination, such as, for example, differential treatment or disadvantage seem more basic to the concept (Thomsen, 2013, p125). People would be confused if someone claimed she was discriminated against, but she was not treated differently from others, or suffered no disadvantage. In contrast, Thomsen suggests that "Few people would … bat an eyelid were I to refer to my being fired because my employer dislikes persons whose first name

---

[14] "We find that competition among workers promotes employers' preference for discrimination. The reason for this is that with discrimination, employers can enhance the expected quality of signals from selected workers when the signals about workers' ability are noisy. In the case of competition, what matters is the quality of winning workers' signals. Discrimination generally increases the signaling incentive of the workers in an advantaged group and thus enhances the quality of their signals, whereas it lowers the incentive of disadvantaged group workers. This implies that discrimination will enhance the overall quality of signals as long as the effect on the advantaged group is significant and their probability of winning increases as a result. This indicates when discrimination is more likely to occur and, if it does occur, who should be the advantaged group. If one group's presence is dominant in the population, workers from the majority group are more likely to be present in any given competition. Then, discrimination in favor of them is likely to enhance their probability of winning and the overall quality of signalling in any competition. Therefore, employers are likely to prefer discrimination that favors the majority group." (Mialon & Yoo, 2014, p1).

[15] Mialon and Yoo (2014) certainly see no problem in referring to the employers' actions in their modelling as 'discrimination'.

starts with an "F" as a case of discrimination" (Thomsen, 2013, p125). I think Thomsen is right, and that people would have no trouble seeing the employer's action as at least potentially wrongful discrimination.

The employer's discrimination against a group of applicants due to the first letter of their first name is unlikely to lead to systemic disadvantage. Hardly anyone will share the strong distaste of the employer in Thomsen's example, and there is no other incentive given to discriminate. By contrast, statistical discrimination is incentivized by a desire to minimize the risk of a bad outcome. Part of the moral concern with statistical discrimination could be that this incentivized desire may lead to statistically 'risky' groups becoming systemically disadvantaged. This could make some of these groups 'socially salient' or perhaps even 'castes' that suffer more widespread social welfare disadvantages. I agree this is a moral concern with statistical discrimination. However, what should we say about the group members who are disadvantaged on their group basis *before* the group achieves social salience? It would seem strange to suggest that they cannot possibly have been wrongfully discriminated against – that this can only apply to those disadvantaged *after* the group attains social salience. Yet this is what theorists focussing on 'social salience' or 'castes' seem committed to maintaining.

Consider, for example, a situation where women on social welfare benefits receive access to some educational initiatives via recommendations from social workers. Suppose that researchers combing through statistics discover a correlation between women who have partners in prison and lower success rates in educational initiatives. Suppose that a social worker reads the research and decides not to offer access to educational initiatives to women whose partners are in prison. Instead, the social worker offers the positions to other women on benefits who would be next in line for the initiative. The group 'Women whose partners are in prison' is not a socially salient group. It is possible that perhaps through this sort of treatment they may *become* a socially salient group; but perhaps not. In any case, it seems unobvious why the social worker's action *now* should not count as (potentially wrongful) discrimination against such women.

In conclusion, differential disadvantageous treatment must be based on some kind of grouping to count as discrimination – the treatment must not be idiosyncratic to the individual affected, such as not hiring someone due to disliking the unusual way they style their hair. However, the group member does not need to be from an existing socially disadvantaged group, or even from a currently socially salient group, for their disadvantageous treatment to count as morally concerning discrimination.

2.2 Rejecting agent-based accounts of discrimination

As we saw in chapter one, philosophers have not reached a consensus on how to specify the moral wrong(s) of discrimination. Philosophers usually agree over what sorts of actions are recognized as wrongful discrimination, while disagreeing over what makes discriminatory actions bad or wrong as such. I see three main categories for accounts of the wrongs of discrimination. *Agent-based accounts*, such as those of Richard Arneson (2006) and Peter Vallentyne (2006), consider the particular moral wrong of discrimination to be located in the false or prejudiced beliefs or hostile attitudes of the agent (the decision-maker). *Consequentialist accounts*, such as those of Lippert-Rasmussen (2006), Moreau (2010) and Segall (2012) consider the moral wrong to be a matter of the consequences for the people affected by discrimination. *Expressivist accounts*, such as those of Hellman (2008) and Eidelson (2013), specify the (or a) moral wrong in the objective meaning expressed by the action: for example, if it demeans people (Hellman, 2008) or fails to properly recognize a person as an individual (Eidelson, 2013).

These categories of accounts of wrongful discrimination do not form an exhaustive list, and some accounts include elements of each category. Larry Alexander (1992), for example, holds that discriminatory actions are morally wrong if they fail to correctly acknowledge a person's moral equality or moral worth in a way that harms people, thus including both an agent-based and consequentialist perspective. Overall, however, I think the categories represent some key ways wrongful discrimination has been conceived in the philosophical literature. I will argue that the agent-based accounts of Arneson and Vallentyne, under which statistical discrimination cannot count as a form of morally wrong discrimination, ought to be rejected. I will then argue that with the exception of Kasper Lippert-

Rasmussen[16], the remaining philosophers ought to take more account of statistical discrimination in the ongoing philosophical discussion of the moral wrong(s) of discrimination.

Richard Arneson (2006) has argued that discrimination shows flawed beliefs or hostile attitudes based on group features. Arneson (2006, p779) defines wrongful discrimination as defective conduct based on "unjustified hostile attitudes toward people perceived to be of a certain kind or faulty beliefs about the characteristics of people of that type". Similarly, Peter Vallentyne (2006) sees discrimination as invidious when it stems from your belief about a socially salient feature, where

> 1) the individual is not morally or prudentially responsible for having that feature; and

> 2) the treatment is based on

>> a) a mistaken belief in the moral inferiority of groups with that feature,

>> b) significantly empirically mistaken beliefs about groups with that feature, or

>> c) hatred of groups with that feature (Vallentyne, 2006, pp982-3)

Both theorists see discrimination as differential treatment of some individuals based in the agent's (the decision-maker's) flawed negative attitudes or beliefs about groups of which the individual is a part: either hostile attitudes toward a group or faulty or mistaken beliefs about the characteristics of a group[17].

---

[16] Lippert-Rasmussen is excepted because he has featured statistical discrimination prominently in his work (Lippert-Rasmussen, 2007; 2011; 2014); I will address his work on statistical discrimination in my next chapter.

[17] The hostile attitudes would be a form of bias, while the faulty beliefs would probably be seen as leading to prejudices. Recall from the explanation early in chapter 1 that bias is a deep-grained hostility against the group as a whole, e.g. misogyny, overt racism of the Ku Klux Klan variety, extreme homophobia. Prejudice is "a faulty or inflexible generalization about members of a group" (Allport 1954, cited in Duell & Landa, 2015, p4).

Recall now that statistical discrimination involves neither false beliefs about a group, nor hostile attitudes from the agent. The agent-based accounts of wrongful discrimination from Arneson and Vallentyne pose the biggest problem for my claim that statistical discrimination ought to feature more prominently in accounts of the moral wrongs of discrimination. This is because they rule out statistical discrimination **by definition** as of no moral concern as a form of discrimination. However, I think the research on stereotypes suggests two problems with agent-based accounts. First, they focus on the agent's beliefs or attitudes towards the groups discriminated against as being at best quite negative, at worst actively hostile. This will cover some forms of discrimination, but only the extreme forms such as misogyny, explicit racism and explicit homophobia. But the research suggests that the attitudes that can result in the discriminatory outcomes we are concerned about, for example, in hiring job applicants or in selecting tenants for rental properties, are not of this nature. It is not that decision-makers are actually hostile toward some groups, but rather that they have a positive preference for other groups.

Suppose that we draw an 'attitudinal line' thus:

_____

extreme hostility   *neutrality*   *extreme sympathy*

Suppose that any attitude to the right of the midpoint is positive, and any attitude to the left of the midpoint is negative. Agent-based accounts (such as those of Arneson and Vallentyne) seem to position the discriminator's attitude well to the left of the midpoint. But Charles Stangor (2009, p2) in his historical overview of the development and current state of research into prejudice reports that "the evaluations of most outgroups are overall positive - at least above a neutral point." The research into discriminations involving ingroups and outgroups, where outgroups represent minority groups that typically suffer discrimination, suggests that people have a less favourable attitude toward outgroups. However, this is a less favourable attitude relative to an ingroup, not relative to a neutral midpoint. People's attitudes are not typically actively hostile toward outgroups; instead, people can have a neutral or mildly positive attitude towards these groups (Stangor, 2009, p2). While people

do have a more negative attitude toward minority groups than they do toward majority groups, it is *not* typically an overall very negative or hostile attitude to minority groups.

Daniel Balliet, Junhui Wu and Carsten DeDreu (2014), for example, conducted a study examining discrimination when different groups are required to cooperate. They separated participants into ingroup members, outgroup members, and unclassified strangers. They found that there was little difference in how ingroup members treated outgroup members as compared with how they treated unclassified strangers. There was a difference however, in how ingroup members treated fellow ingroup members, relative to both other groups – ingroup members favoured each other. Balliet, Wu and DeDreu suggest their results support the conclusion that favouritism toward the ingroup is what drives intergroup discrimination in a cooperative setting (Balliet, Wu & DeDreu, 2014, p19).

This suggests that discrimination can happen because people have a stronger positive preference for ingroups, rather than a neutral attitude to ingroups and a negative or hostile attitude toward outgroups. Of course, this does not change anything about how much discrimination occurs – it is not suggesting there is any less discrimination by majority members against minority group members. If we are focused on consequences, it makes no difference whether these outcomes are motivated by hostile attitudes or a positive preference for ingroups. But this difference in the conception of the discrimination does make a difference to what we should accept in a theory of wrongful discrimination. It suggests that a theory that relies on hostile attitudes may only cover a small proportion of the discrimination that occurs today and that is the subject of discussion in the literature.

The second problem with these accounts of discrimination concerns the claims of significantly mistaken or faulty beliefs about the characteristics of groups. This claim is essentially about the accuracy of stereotypes. Stereotypes are cognitive structures[18]. Lee Jussim and colleagues cite Ashmore and Del Boca's definition of a stereotype as "a set of beliefs about the personal attributes of a social group" (Jussim et al, 2009, p201). Hamilton and colleagues (2009, p9) add beliefs about characteristics and behavior patterns associated with a particular group. They note that "Once formed, that set of beliefs is applied to all members of the group, generalizing across individuals, despite the fact that those persons

---

[18] schemas, prototypes or exemplars (Stangor, 2009, p3).

may show considerable variation in numerous respects." (Hamilton et al, 2009, p180). An example of a stereotype would be the belief that older workers have trouble learning new things.

The problem with the claim of significantly mistaken or faulty beliefs about the characteristics of groups is that it is not clear that this is always, or even often, true. In a few areas, such as stereotypes of groups based on nationality, people's stereotypes have been shown to be quite inaccurate; but this is not the case for all stereotyping. A group of researchers (Lee Jussim, Thomas Cain, Jarret Crawford, Kent Harber and Florette Cohen) has been reviewing the (admittedly few) empirical studies of racial or ethnic and gender stereotypes that allow the comparison of the perceivers' stereotypes against the actual data for the groups, in order to assess the stereotypes for accuracy. These studies suggest that often people's stereotypes of groups are reasonably accurate, and certainly not wildly inaccurate.

For racial or ethnic stereotypes, the researchers looked at four studies, which covered beliefs about group differences on achievement, athleticism, criminality, and demographic factors such as having a high school degree or being on welfare. For gender stereotypes, the researchers looked at eight studies. The studies covered beliefs about sex distribution into different occupations, sex distribution into different majors and grade point averages, sex differences in nonverbal behavior, and beliefs about the attitudes of men and women. The researchers found that most stereotypes are pretty accurate at the group-level (Jussim et al, 2009). By 'pretty accurate', the researchers mean either within 10% of what the actual statistics show, which they class as 'accurate'; or within 20% of what the actual statistics for the group show, which they class as a 'near-miss' (as opposed to inaccurate) (Jussim et al, 2009, p.204). (The studies were all from the US or Canada, and the majority used undergraduate students; for the studies that used adults, the adults' stereotyping was more accurate).

The findings from these studies suggest that we should be hesitant to characterize discrimination as stemming from false beliefs or faulty beliefs on the part of the agent about characteristics of groups. Overall, we have seen that agents who discriminate need not, and often will not, have hostile attitudes towards the group discriminated against. We have also

seen that such agents may not have flawed beliefs about the characteristics of the group against which they discriminate either. On these grounds, I suggest that positions on wrongful discrimination that are based purely in the falsity of an agent's beliefs about a group or hostile attitudes toward a group ought to be rejected as inadequate to account for the phenomena of discrimination.

2.3 Accounts of discrimination that cover statistical discrimination

2.3.1 Stereotype and stigma

Unlike the purely agent-based accounts of wrongful discrimination, the accounts of both Shlomi Segall (2012) and Sophia Moreau (2010) do not preclude statistical discrimination. While neither theorist focusses on statistical discrimination, their accounts of morally concerning discrimination can accommodate statistical discrimination, at least *prima facie*. I think the accounts of Segall and Moreau suggest a view of wrongful discrimination as connected to the effects of stigmas for certain groups of people. In what follows I will briefly outline the positions of Segall and Moreau, explain the concept of stigma, and note why I think their views align with a stigma-based view of discrimination. I will then argue that stereotyping, rather than stigma, is active in differential treatment that excludes some people; and that this gives us reason to focus more closely on statistical discrimination.

Shlomi Segall (2012) argues that what is bad about discrimination is its undermining of equality of opportunity, broadly conceived. Segall notes (2012, p83) that his concept of equality of opportunity is wider than lacking the same opportunity as someone else applying for a job, for example. It is more like equality of opportunity across all the aspects of society that can impact on our welfare, such as, for example, housing or healthcare access. So Segall sees discrimination as bad when it comprises actions or policies that constrain an individual's equal opportunity with regard to social welfare. For Segall, if statistical discrimination undermines our equality of opportunity across different aspects of society that can impact on our welfare, then it should be seen as bad.

Whereas Segall sees the badness of discrimination in terms of treatment that differentially constrains our objective welfare opportunities, Sophia Moreau is concerned about differential treatment that constrains our subjective autonomy. Moreau (2010) defines

51

wrongful discrimination as differential treatment that affects our 'deliberative freedoms', our liberties to access core life opportunities without worrying about negative judgements on our 'normatively extraneous' traits such as sex or race or nationality. Moreau (2013, p136) thinks that who we are is, at least in part, constituted by the choices we make. Discrimination unfairly constrains our choices through other people's assumptions about our extraneous traits and through their treatment of us. Her account paints wrongful discrimination as the making of distinctions between socially salient groups that can constrain the autonomy of groups that are discriminated against. If statistical discrimination ends up negatively constraining people's deliberative freedoms and autonomy, then Moreau would see it as wrong.

Hatzenbuehler, Phelan and Link (2013) define stigma as "the co-occurrence of labeling, stereotyping, separation, status loss, and discrimination in a context in which power is exercised" (Hatzenbuehler et al, 2013, p2). Their definition focusses on identifying the inputs that combine to create stigma. In terms of the output of this combination, Erving Goffman (2009) conceives of stigma as a spoiled social identity, when a person is "… reduced in our minds from a whole and usual person to a tainted, discounted one" (Goffman, 2009, p3). A stigma is a negative association, indicating a person is of a "less desirable kind" (Goffman, 2009, p3). Goffman (2009, p4) distinguishes three types of stigma: perceived bodily deformities, perceived character blemishes (which may be generated from a known history of mental disorder, addiction, unemployment etc) and "tribal stigmas" such as race, nation, and religion. Stafford and Scott (1986, p77) note that stigmas can attach to many group features including criminality, disability, homosexuality, alcoholism, being black, receiving welfare, illiteracy, mental illness, and obesity, amongst others.

Katz (2014, p2) suggests that "…. there are important attributes that almost everywhere in our society are discrediting." What aligns the views of Segall and Moreau with the concept of stigma is their linking of wrongful discrimination with group features that will have a persistently negative effect across the social contexts in which a person lives their life, rather than, say, sometimes a negative impact for them but sometimes a positive impact. Moreau also has a concern with the psychological impact of discrimination – the impact on people's freedom to deliberate about their core life opportunities without the concern that some of their traits will be held against them. Stuber, Meyer and Link (2008) note the

psychological impact of stigma: stigmas can become internalized in group members and cause stress. This can occur even if people can hide a stigmatising feature. John Pachankis (2007) writes of the stress and challenges facing people with a concealable stigma, such as epilepsy, illiteracy, infertility, deafness, or being from a working-class background.

So, a stigma is a negatively-valenced feature of a group that can affect impressions of and actions toward group members. Stereotypes, however, while often negative, can also be positive. For example, the stereotype that Asian students are good at mathematics is a positive stereotype. Also, stereotypes can be relatively specific; for example, the stereotype in New Zealand that Asian drivers have poorer driving skills is salient only in the context of driving. However, when there is a stigma attached to a group membership, we would expect the group-identity to cast a negative pall over social interactions in many areas where it was salient. We would not expect a stigmatized feature to work in favour of individuals categorized under the feature, relative to individuals without that feature. These aspects of stigma, I think, align it with the concerns for what is wrong or bad about discrimination expressed by the main focus of each of Segall and Moreau.

If discrimination works on the level of stigma, then belonging to more than one stigmatized group should magnify discrimination in important contexts. Suppose that being Black and being homosexual are stigmatized attributes in the United States – each gives rise to discrimination. This would indicate that if you are applying for a job and you are Black AND gay you are doubly disadvantaged, and you will face heightened discrimination. However, evidence from research on employment suggests that combining group statuses that are individually associated with disadvantage or discrimination does not necessarily compound discrimination in this way. For example, David Pedulla (2014) from Princeton carried out a study on the intersection of sexual orientation and race in employment. His research suggests that in the US context gay Black men are likely to face less discrimination in employment than either heterosexual Black men or gay white men. He theorized this is because in the United States, there is a stereotype of Black men as likely to be more aggressive than White men, and gay men as more likely to be effeminate – which translated as less aggressive - than heterosexual men. Assuming that strong aggression and strong effeminacy in males are both features unattractive to employers, the combination of these opposing stereotypes would seem to cancel or work against each stereotype for gay Black

men. This is hypothesised to *reduce* the stigma for gay Black men relative to either gay White men or straight Black men. In this example, it seems that the stereotypes are driving the resulting discrimination, rather than the general negative stigma attached to each of the stereotypes.

Stigmas attached to race or minority sexual orientation are well-documented in the literature (see, e.g., Hatzenbuehler, Phelan & Link, 2013). Other stigmas that are less well-known may interact with these. One stigma that affects women of childbearing age seeking employment, known as the 'motherhood' penalty, is not covered in the literature on stigma, but instead in general sociology literature (see e.g. Budig & England, 2001; Correll, Benard & Paik, 2007). The 'motherhood' penalty is where employers presume a woman of childbearing age will be likely to be less productive as an employee due to taking more time off or more time out of their employment for parenting, and so the rate of hiring or pay for these women is lower than for other applicants. Research by Alyssa Schneebaum (2013, using US data and furthering research by Amanda Baumle, 2009) confirms that motherhood is typically negatively correlated with wage rates for heterosexual working women; however, motherhood is positively related to wage rates for lesbian working women. The hypothesis offered here is that employers tend to categorise lesbian working mothers as similar to working fathers (who are presumed to have a wife to absorb the work-impacting costs of children and domestic care arrangements), and so lesbian working mothers are not subject to the 'motherhood' penalty that afflicts heterosexual working women.

Even within the group of working women who are mothers, people engage perceptual stereotypes that can counteract stigmatized identities. For example, Rebecca Glauber (2007) studied the interaction of motherhood and race, and found that the 'motherhood penalty' applied mainly to white women, not Hispanic or Black women. Kathleen Denny (n.d.) has also looked at workplace evaluations of parents by race. She too found that the perceptual motherhood penalty seemed to be strongest for white women - stronger for than for Black or Asian mothers (Denny, n.d, p26). Denny hypothesizes there is a stereotype of white mothers as highly involved in their children's lives and so less committed to the workplace, giving rise to a 'motherhood' penalty. It seems that combining group statuses that are individually associated with discrimination, such as family status, sexual orientation and race, does not necessarily compound the discrimination – it can lessen it. Again, it

seems that stereotypes are motivating the resulting discriminations, and that the interacting stereotypes can reduce the impact of a general stigma attached to a group.

Overall the evidence indicates that context is important for how discriminations will impact on equality of welfare opportunity and whether (or the extent to which) stigmatised traits will count as costs against a person. This finding, I suggest, is not intuitive. I think that intuitively we would expect that if Simon sees Beatrix has a stigmatized feature, Simon should perceive this as a negative mark in context A, B or C, and Beatrix having an extra stigmatized feature should just make things worse. This would fit with Segall's and Moreau's views on the moral wrongnesses of discrimination. The evidence that stereotypes can counteract stigma suggests to me that stereotypes are the conceptual drivers that result in differential treatment, rather than stigma per se. This would indicate that *stereotyping* is the active ingredient in stigmatized identities as regards discrimination. Statistical discrimination, with its potential for systemic disadvantage due to the rational use of stereotypes, should thus be a fruitful phenomenon to consider.

2.3.2 Can statistical discrimination be demeaning?
Like the accounts of Segall and Moreau, the expressivist positions on wrongful discrimination from Deborah Hellman (2008) and Benjamin Eidelson (2013) can accommodate statistical discrimination, and both Hellman and Eidelson include some focus on statistical discrimination. However, Hellman's discussion of statistical discrimination seems to dismiss the likelihood that it will be wrong as a form of discrimination. I will argue that Hellman's focus on the accuracy or inaccuracy of the statistical generalization is too narrow to capture what may be wrong with discrimination on the basis of group risk statistics. I will explain how Eidelson's position does better on this, and draw on some research to support this view.

Deborah Hellman (2008) argues that the moral wrong in discrimination is not about the beliefs or attitudes of the agent, nor about the consequences of the discrimination. Instead, the moral wrong lies in what the discriminatory act **expresses**. Morally wrongful discrimination involves drawing distinctions between people in a way that **demeans** a person, which means "to treat another as not fully human or not of equal moral worth"

(Hellman, 2008, p36). The social context gives the action its demeaning quality, expressing denigration in a way that has the capacity to put someone down. As noted in chapter 1, it seems prima facie unlikely that statistical discrimination demeans people, as it does not obviously express anything about the moral worth of the people involved. On Hellman's definition, statistical discrimination does not appear likely to be morally wrong.

Hellman (2008) addresses statistical discrimination in an example in her book's fifth chapter, which discusses the importance (or perhaps lack thereof) of accuracy and irrationality to determining discrimination. She wants to argue, not that inaccurate classifications are never morally problematic, but rather "...that their accuracy or inaccuracy is not why they run afoul of the requirement to treat people as moral equals" (p121). Hellman's method for testing her claim is to look at examples that do not seem to involve other wrong-making features. Her example is of a university that believes math and science majors are poorer swimmers than other majors and requires math and science majors to pass a remedial swimming course. Hellman takes a pragmatic view of this discrimination, noting that generalizations may be unavoidable, and considers whether institutions that do not make efforts to alleviate burdens imposed by false generalizations are engaging in morally wrongful discrimination. She suggests that in order to see whether this is a problem for treating people as being of equal moral worth, we can look at what the alternatives are.

The alternatives Hellman suggests (2008, p123) are that the institution uses an accurate proxy, or that it treats everyone the same (treating everyone the same in this case would be to make everyone or no one take the swimming course). She notes the efficiency argument against using a more accurate proxy: either it may cost more to test for, or it may cost more to discover what the more accurate proxy is. Hellman then asks about generalizations whose accuracy is easy to test (at little cost), but where the institution continues to use a poor proxy. In this case she suggests we say that burdened persons are on the receiving end of a *blunder*. She notes that we cannot expect blunder-free policies and that this is unfortunate but no violation of rights, especially the right to be treated as equally worthy of moral concern. She suggests the persons concerned ought to complain and those involved should seek to rectify the decision-making process that allowed the error.

I think Hellman is right that that the accuracy of a proxy is not in itself a deciding factor in judging the ethics of statistical discrimination. However, I think Hellman does not take enough account of a more fundamental characteristic that is typical of statistical discrimination: that of judging individuals solely on the basis of group risk. For any person, 'group risk' is a degrading lens through which to be viewed. The 'group' facet strips us of our individualized features; it ignores our identity as an individual. The 'risk' facet presents us as posing a threat that needs to be mitigated. Thinking back to the example of the university and the swimming lessons, suppose that the university chooses to continue using a poor proxy because an individualized test is more expensive. An institution that does this is making a unilateral decision to disadvantage some students on the basis of a group risk characteristic. If an institution does this, I think the institution could be demeaning those students.

Hellman (2008, p23) is clear that wrongful discrimination is not (or not solely) about discrimination against a group; the unfairness is (or is also) about the treatment of the individual discriminated against. Recall Hellman's initial definition (2008, p36) that to demean someone is "to treat another as not fully human or not of equal moral worth", and that the social context gives the action its demeaning quality. If statistical discrimination on the basis of group risk does not take account of someone's individual identity, in a social context where this is important, then perhaps it does not treat a person as fully human. Not all social contexts are important to our individual identity. It could be that where what is at stake is relatively unimportant to our individual identity, we will not count statistical discrimination as demeaning. Perhaps being asked to do some swim lessons is unimportant in this way. However, other contexts, such as what we do for a living or how we shape ourselves through our choice of subjects to study, might be more important. This suggests a need to focus on more substantive instances of statistical discrimination, such as those relating to goods that are important to a person's individual identity.

The claim that discrimination is morally problematic because of failing to treat someone as an individual person seems intuitive, even if it has received little support in the philosophical

literature[19]. Benjamin Eidelson argues that we should see the failure to recognise individuality as part of the badness of discrimination due to the importance of individuality to our autonomy[20]. To respect a person as an individual we must satisfy two requirements, according to Eidelson (2013, p375): a *character* requirement and an *agency* requirement. The character requirement is that we must give reasonable consideration to the way a person has used her autonomy in making her decisions over her life. The agency requirement is that "we must not make predictions about her choices in ways that demean the role of her autonomous agency in making up her own mind" (Eidelson, 2013, p57). Discrimination against a person due to a group feature fails "to treat him as in part a product of his own past efforts at self-creation, and as an autonomous agent whose future choices are his own to make"(Eidelson, 2013, p395).

Eidelson seems to suggest that the morally problematic aspect of discrimination's failure to treat a person as an individual is expressive: it is about actions that do not properly acknowledge someone's dignity as a person, do not accord them proper moral recognition. Eidelson's concern would suggest discrimination that excludes or disadvantages applicants purely on the basis of group risk statistics is morally concerning. He discusses the example of an employer simply assuming a female applicant is likely to be less productive because of a decision to have children. Eidelson would note that in making this assumption, the employer gives no particular consideration to the way the applicant has used her autonomy in making decisions over her life. Moreover, the employer fails to treat her as an individual if he forms judgements "in a way that treats her as determined by statistical tendencies" (2013, p381), rather than allowing some acknowledgement that she is an autonomous agent who can make her own choices. Statistical discrimination would seem to be a paradigm case of failing to treat people as individuals.

The concern Eidelson raises with discrimination squarely characterizes what I think is the primary non-consequentialist concern with statistical discrimination; that is, that assessing applicants based on the statistical characteristics of some groups they belong to takes little

---

[19] Ferdinand Schoeman (1987), Frederick Schauer (2003) and Kasper Lippert-Rasmussen (2011) argue against it; I will address these arguments in my next chapter.
[20] Note that Eidelson says that the concern he raises is only a partial and prima facie wrong-making aspect of a discriminatory action. This means that other aspects of discriminatory actions such as good consequences or fair distribution, could override the wrongness.

account of the individual identity of applicants. Psychological research reveals that the very process of generalizing across individuals leads to the perception of homogeneity among group members (Hamilton et al, 2009, p179); that is, the very process of stereotyping leads decision-makers to see the stereotyped group as more homogenous. This means that decision-makers can see group members as more likely to resemble each other on various traits and to be more similar to each other in general (Hamilton et al, 2009, p182). This suggests that the process of applying a stereotype to an individual, no matter whether it is a positive or negative stereotype, can work at a psychological level to further de-individualise the group member from the perspective of the decision-maker.

If we treat a person differently just because she belongs to a particular group, we may not make an effort to take account of the ways she has exercised choices that differentiate her from the group. Furthermore, we also decide which feature of her - that feature of belonging to that group - will loom large in our assessment of her. As regards treating someone as an individual, these are flaws, and applying statistical discrimination to exclude applicants to organisations exhibits these flaws. In chapter three I will argue that the moral problem with this discrimination is not the accurate or inaccurate identification of a feature as characteristic of a group; the problem is the use of a group generalisation to judge an individual applicant in place of the use of individualizing evidence.

Conclusion

In this chapter I have argued that the moral salience of discrimination based on group risk statistics is under-recognised. My first approach was to argue against concepts of discrimination and normative accounts of discrimination which seem either to rule out or seriously marginalize the potential wrongness of this statistical discrimination. My second approach argues for the relevance of this statistical discrimination to accounts of discrimination that can accommodate it. Overall, I have argued that statistical discrimination raises an important moral issue. In conclusion, my analysis suggests that a key part of morally concerning discrimination is simply treating individuals on the basis of group risk characteristics in place of using individualized information. My next chapter will critically examine this conclusion. I will address objections put forward by Ferdinand Schoeman (1997), Frederick Schauer (2003) and Kasper Lippert-Rasmussen (2007, 2011, 2014). I will

discuss the connection of statistical information to individuals, illustrate how this differs from that of individualized evidence, and offer views on what is valuable about treating a person as an individual.

Chapter Two References

Arneson, Richard J. (2006) What is wrongful discrimination? *San Diego Law Review,* 43: 775-808.

Balliet, Daniel, Wu, Junhui and De Dreu, Carsten K W. (2014) Ingroup favouritism in cooperation: A meta-analysis. *Psychological Bulletin*. Doi:10.1037/a0037737

Baumle, Amanda K. (2009) The demography of sexuality and the labor market*. International Handbook on the Demography of Sexuality.* 5, (Pp243-256).

Budig, Michelle J and England, Paula (2001) The Wage Penalty for Motherhood. *American Sociological Review, 66*(2): 204-225. Retrieved September 2016 from http://www.jstor.org/stable/2657415

Correll, Shelley, Benard, Stephen and Paik, In (2007) Getting a job: is there a motherhood penalty? *American Journal of Sociology*, 112 (5): 1297-1339.

Denny, Kathleen E. (n.d.) Workplace evaluations of parents by race: Unraveling perceptual penalties and premiums. Retrieved July 2016 from http://www.tessexperiments.org/data/Denny366.pdf

Duell, Dominik and Landa, Dimitri (30 March, 2015) Determinants of discrimination in a strategic setting. Working paper. Retrieved September 2016 from http://dominikduell.com/papers/ATT_paper.pdf

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York: Routledge.

Eidelson, Benjamin (2013) Treating people as individuals. In Deborah Hellman and Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp354-395). Oxford: Oxford University Press.

Glauber, Rebecca (2007) Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. *Journal of Marriage and Family*, 69: 951–961. doi:10.1111/j.1741-3737.2007.00423.x

Hamilton, D. L., Sherman, S. J., Crump, S. A., & Spencer-Rodgers, J. (2009) The role of entitativity in stereotyping: Processes and parameters. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 179-198). Philadelphia, PA: Psychology Press.

Hatzenbuehler, Mark L, Phelan, Jo C and Link, Bruce G. (2013) Stigma as a fundamental cause of population health inequalities. *American Journal of Public Health, 103* (5): 813-821. Retrieved August 2016 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3682466/

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, MA: Harvard University Press.

Jussim, Lee, Cain, Thomas R, Crawford, Jarret T, Harber, Kent, Cohen, Florette (2009) The unbearable accuracy of stereotypes. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 179-198). Philadelphia, PA: Psychology Press.

Katz, Irwin (2014) *Stigma: A social psychological analysis*. New York: Psychology Press.

Lippert-Rasmussen, Kasper (2006) Private discrimination: A prioritarian, desert-accommodating account. *San Diego Law Review,* 43: 817-856.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *The Journal of Political Philosophy*, 15: 385-403.

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15: 47-59.

Lippert-Rasmussen, Kasper (2014) Statistical discrimination. Ch. 3 in *Born free and equal? A philosophical inquiry into the nature of discrimination*. Oxford University Press Online. DOI: 10.1093/acprof:oso/9780199796113.001.0001

Mialon, Sue H and Yoo, Seung Han (2014) Employers' preference for discrimination. Emory Economics Working Paper Series 1402. Retrieved from http://economics.emory.edu/home/assets/workingpapers/smialon_14_02_paper.pdf

Moreau, Sophia (2010) Discrimination as Negligence, *Canadian Journal of Philosophy*, 40:sup1: 123-149.

Moreau, Sophia (2013) In Defense of a Liberty-based Account of Discrimination. In Deborah Hellman and Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp.71-86). Oxford: Oxford University Press.

Pachankis, John E. (2007) The psychological implications of concealing a stigma: A cognitive-affective behavioral model. *Psychological Bulletin*, 133 (2): 328-345.

Pedulla, David (2014) The Positive Consequences of Negative Stereotypes: Race, Sexual Orientation, and the Job Application Process. *Social Psychology Quarterly*, 77 (1). doi: 10.1177/0190272513506229

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Schneebaum, Alyssa (2013) Motherhood and the lesbian wage premium. University of Amherst-Massachusetts Working paper No. 4. Retrieved from *https://www.umass.edu/economics/publications/2013-04.pdf*

Schoeman, Ferdinand (1987) Statistical vs direct evidence. *Nous*, 21 (2): 179-198.

Segall, Shlomi (2012) What's so bad about discrimination? *Utilitas*, *24*(1) : 82-100. doi:10.1017/S0953820811000379

Stafford, Mark C and Scott, Richard, R. (1986) Stigma, Deviance and Social Control. The Dilemma of Difference, *Perspectives in Social Psychology*, pp77-91.

Stangor, Charles (2009) The study of stereotyping, prejudice, and discrimination in social psychology: A quick history of theory and research. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp3-22). Philadelphia, PA: Psychology Press.

Stuber, Jennifer, Meyer, Ilan and Link, Bruce (2008) Stigma, prejudice, discrimination and health. *Social Science & Medicine*, 67 (3): 351-357.

Thomsen, Frej Klem (2013) But some groups are more equal than others: A critical review of the group criterion in the concept of discrimination. *Social Theory and Practice*, 39 (1): 120-146.

Vallentyne, Peter (2006). Left libertarianism and private discrimination. *San Diego Law Review, 43*: 981-994.

Chapter three   Statistical discrimination, treating applicants as individuals, and agency[21]

Introduction

When individuals apply to an organisation to access a good or service offered by the organisation, some person or group has to assess the application and make a decision on it. Decision-makers can be interested in predicting the likelihood that an applicant will impose a significant cost on the organisation or on other people. For example, parole boards consider information relevant to predicting the likelihood that a prisoner will re-offend once released. In the commercial context, insurance companies calculate the risk that an insured will make a claim, and managers want to ensure they hire a productive and reliable employee. Even in the education field, institutions are interested in predicting which applicants for tertiary study are a higher risk for dropping out or failing.

To assess applicants, decision-makers can use evidence we might think of as 'individualised'. Individualised evidence is evidence of a person's particular combination of skills or competencies and character and actions that marks them out as an individual. This could include the applicant or their referees attesting to their attributes and personal history, through character or academic references. Individualised evidence could also include an applicant's efforts, achievements and competencies, as shown through, for example, rehabilitation and anti-violence programs completed by prisoners, work tests by job applicants, or healthy exercise regimes by insurance applicants. The combination of character attributes, actions and skills that is particular to an applicant is what distinguishes them as an individual.

Conversely, instead of taking into account 'individualised' evidence, decision-makers can identify a feature or features correlated with risks of undesired outcomes across the group of people that have that feature, and assess applicants based on these risk statistics. When decision-makers use features correlated with group risk to exclude 'risky' applicants, this is a form of statistical discrimination. The particular features will differ depending on the

---

[21] The first two paragraphs of the introduction, some of the section on Schoeman and most of the section on Schauer form part of my book chapter "Risky business – the ethics of judging individuals based on group statistics", published in August 2016. See Scholes (2016).

purposes of the organisation and the relevant laws, but could include factors such as age, gender, ethnicity, area of residence, employment status and recreational activities. These features are attributes of the individual, but the decision-maker is not using them as individualized evidence. This is because the feature or features, used in this way, do not distinguish the individual. All people are men, women or intersex; millions of people are any particular age; thousands live in a particular area and are employed. I will argue that the feature or features, used in this way, are not really about the individual. A decision-maker who isolates some features that are statistically correlated with higher risk and uses these to downgrade or exclude an applicant is bypassing the applicant's individuality.

In the philosophical literature, statistical discrimination based on group risk is addressed explicitly and in some detail by Stephen Maitzen (1991)[22]; Amanda Baumle and Mark Fossett, (2005); and Deborah Hellman, (2008, as discussed in the previous chapter). However, the most detailed and influential treatments have been those of Ferdinand Schoeman (1987), Frederick Schauer (2003) and Kasper Lippert-Rasmussen (2007, 2011, 2014). These three philosophers have presented arguments that challenge the distinction between using group-based evidence and individualized evidence when assessing people. I think the distinction still has some value, however, and I will attempt to defend this view in responding to their arguments.

I think that the concept of treating a person as an individual is morally important, even if it is difficult to pin down. As my focus is the organisational selection of applicants, I start with some analysis of what it is to assess an applicant as an individual. I then address arguments from Schoeman, Schauer and Lippert-Rasmussen and try to highlight what is missing from assessments that use purely statistical evidence. I conclude with a section expressing my view on what is valuable about treating a person as an individual, supported by points from

---

[22] Stephen Maitzen (1991) argues policy makers should use social cost-benefit analysis to determine the appropriateness of statistical discrimination in effecting good or permissible social policies. A social policy involving statistical discrimination is justified if it produces an overall positive outcome (p.30). He suggests that the level of statistical discrimination used in a social policy "is justified if and only if the information-cost of further statistical refinement equals or exceeds the net social utility to be gained by such refinement" (Maitzen, 1991, p.26). Any use of statistical discrimination is thus morally justified simply if it produces equal or more social utility than any alternative. Maitzen (1991, pp39-40) suggests concepts of rights can be accommodated in his theory by assigning a high negative utility to outcomes that violate our idea of a right; for example, to a policy that would violate our idea of a right to privacy.

David Edmonds (2006) and Benjamin Eidelson (2013). I suggest that trying to treat people as individuals is important (at least in part) because it allows us to prioritise recognising a person as a locus of unique moral agency. This is forestalled by organisational decision-makers excluding or disadvantaging people purely on the basis of a couple of risk statistics for groups of which they are members, so this is a reason to disprefer this sort of statistical discrimination.

3.1 What does treating an applicant as an individual involve?

What does assessing people *as individuals* involve? What amount of information is needed? What if an assessment of an individual has some false information? Have we still assessed the particular individual, just wrongly - or have we not actually assessed the particular individual? As noted in my methodology in chapter one, I aim to take a pragmatic approach in this thesis. As such, I am less interested in these questions in the abstract, and more interested in their implications in an applied setting. Assessments of individuals do not drop out of thin air, they are done by agents who can do a better or worse job of such assessments. I suggest a pragmatic way of answering questions about assessing people as individuals is to focus on the agents who will be doing the assessments, and their context.

In this thesis, the agents are organisational decision-makers who bear some responsibility for these assessments and will take action (or will prompt others to take action) based on these assessments. This is important because it distinguishes these agents from others, such as bystanders, who might use statistical evidence simply for their own speculation. In the latter case, where nothing is likely to come of the speculation, we might not care much about its accuracy. However, in the case of agents who will be deciding whether to exclude or retain applications for organisational goods, I think we can specify that the decision-makers should make a conscientious effort to accurately assess information about individuals. This will mean scanning a reasonable amount of information provided by the individual. However, it is possible an applicant may provide information clearly showing their unsuitability along with plenty of irrelevant information. Decision-makers have time constraints so it does not seem reasonable that decision-makers must fully absorb all

information provided by an applicant. Also, we cannot expect the organisational decision-makers to be omniscient. We must expect that they will make conscientious efforts to assess individuals accurately and still get some things wrong. So I will presume it is possible for an assessment of an individual to contain *some* false information and still be an assessment of that individual.

While a decision-maker's assessment may bear some level of false information and still count as an assessment of that individual, obviously too much false information – or, in the case where very little information is used, too little true information - will mean a person has not been assessed as an individual. I think the assessment must be sufficient to distinguish the person in question from other people in the relevant population. Context matters for identifying the relevant population and what should count as a sufficient distinction. The decision-makers are assessing a person applying to their organisation for a good; the relevant population is the other applicants. Presuming the organisational decision-maker is a stranger to the applicant, the standard for individual assessment must be set lower than the level of knowledge and ability to distinguish the applicant that we might demand of the applicant's close friends or family. I suggest that if the information used to assess an applicant is not mostly true, it is unlikely the person can be treated as the individual they are. However, it is more important that the information relevant to specifying the applicant's qualities concerning the organisational good (job, admission to tertiary education) be true. If an organisational decision-maker assessed a person's skills, work habits and experience correctly, but misread other information such as their birth month or the school they attended, the person could still have been assessed as an individual applicant.

In what follows, I will canvas some possible ways of judging a decision-maker's assessment of an applicant. An assessment of an applicant is a judgement of the applicant's suitability regarding the organisational good. I will start with a low standard, and raise it if it does not seem to distinguish the applicant sufficiently from others. Suppose that our lowest standard is:

Standard 1. Given the reasonably accessible evidence, a decision-maker has assessed applicant X as an individual, even if the assessment turns out to be in

some ways false, if her assessment of applicant X uses information *more likely to be true than false*.

This will not do. There are statements that are more likely to be true than false that apply to many people, and so could have been made even if applicant X were replaced with any of these people. For example, it is more likely to be true than false that applicant X is alive, that applicant X breathes air, that applicant X can move themselves around, and so on. A claim or combination of claims at this level does not count as assessing an applicant as an individual, even if the assessment is more likely to be true than false.

Let us raise the standard to:

Standard 2. Given the reasonably accessible evidence, a decision-maker has assessed applicant X as an individual, even if the assessment turns out to be in some ways false, if her assessment of applicant X uses information more likely to be true of applicant X than of another person randomly selected.

This will not do. There are statements that are more likely to be true than false of any applicant for the opportunity but not any random person, so this does not distinguish applicant X from other applicants. For example, it is more likely to be true of any applicant than of a random person that the applicant knows about the benefit or opportunity, has some knowledge of the organisation and has some evidence regarding themselves that they can provide that is relevant to the benefit or opportunity. These points are more likely to be true than false of any applicant for the opportunity, not just applicant X in particular. A claim or combination of claims at this level does not count as assessing an applicant as an individual, even if it uses information more likely to be true of applicant X than a person selected at random.

Let us raise the standard again:

Standard 3. Given the reasonably accessible evidence, a decision-maker has assessed applicant X as an individual, even if the assessment turns out to be in some ways false, if her assessment of applicant X uses information more likely to be true of applicant X than of another applicant.

This is promising, because it restricts the reference group to the relevant population: applicants. However, this does not work if the other applicant referred to by "another

70

applicant" is quite peculiar. There are statements that are more likely to be true of applicant X than a peculiar applicant, but may equally be true of several other non-peculiar applicants for the opportunity. For example, suppose that there are fifteen applicants for a professional position, a core aspect of which involves verbally explaining details to clients and reassuring clients. Most applicants are at least minimally suitable for the position. Applicant fifteen, however, has no relevant professional or academic qualification, and has a tremendous stutter and permanently blank facial expression due to an amateur tattooist hitting a nerve when doing her full-face tattoo. She also has no experience in the profession, although she has written poetry about the profession for quite a few years – indeed, she wrote this poetry for most of the years that she was doing time. A brief assessment of applicant X may well use information more likely to be true of applicant X than applicant fifteen, yet may not distinguish applicant X from some of the other thirteen applicants.

Let us raise the standard further:

> **Standard 4. Given the reasonably accessible evidence, a decision-maker has assessed applicant X as an individual, even if the assessment turns out to be in some ways false, if her assessment of applicant X uses information more likely to be true of applicant X than of any other applicant.**

This seems to work, as it meets the two grounds I think we want for an organisational decision-maker to assess an applicant as an individual. 1) It is sufficient to distinguish the person in question, applicant X, from other people in the relevant population – all the other applicants; and 2) it does not (at least, not obviously) require the level of knowledge and ability to distinguish applicant X that we might demand of X's close friends or family.

Note that in endorsing standard 4, I intend to make a normative claim about treating an applicant as an individual, but not a moral claim. In other words, I am suggesting this ought to be the standard if we want to judge whether an organisational decision-maker has assessed someone as an individual, but I am not suggesting an organisational decision-maker necessarily ought to do this or is morally good if they do this. My focus at this point is identifying what counts as individual assessment; I will bring in considerations of the moral importance of treating persons as individuals later in the chapter. Obviously, there may be

other aspects of an organisational decision that are more important, morally speaking, than whether an applicant is assessed as an individual.

3.2 Is statistical information about individuals?

Schoeman (1987) is interested in information as evidence, where evidence is thought of as 'direct', such as in an eyewitness report, or instead based on statistics about groups. Schoeman (1987) asks whether we can say we get information about individuals when all we have to go on is statistical classes. 'Direct' information such as from an eyewitness report seems to apply to an individual in a way that statistical information does not. He suggests that perhaps our dislike of using group statistics when judging an individual is due to a perception that the information used is not really *about* the individual concerned. Schoeman (1987, p183) characterises the concern thus:

> [perhaps] 'direct' evidence points to an individuated claim in the sense that it refers explicitly to a specific agent in a specific situation having done a specific thing whereas in the statistical case what the evidence points to is the conjunction of [reference] classes – a generic conclusion.

I propose the following brief scenario to illustrate the concern. Suppose that Jane Comb has become comatose through eating poisoned fast food in New Buckshire. Suppose too that eyewitness accounts of events have a reliability rating of 90%. Compare an eyewitness report of Jane Comb entering an Organics-to-go cafe on the morning of the day she went into the coma, with a report that says 90% of all fast-food servings in New Buckshire are served in Organics-to-go cafes. The former 'direct' evidence refers to a specific agent in a specific situation having done a specific thing. The latter 'statistical' evidence suggests that for any generic person ordering at a fast-food outlet in New Buckshire, there's a very high probability they are ordering at an Organics-to-go cafe. The combination of the reference classes 'person dining out on fast food' and 'New Buckshire fast-food outlets' is evidence of a high likelihood of eating at an Organics-to-go cafe. However, it does not seem to be about Jane in the way that the eyewitness report seems to be about Jane.

Schoeman suggests (p184) there is a problem with dismissing the sort of statistical evidence I have outlined above. The problem is that it would prevent any group-level information from being relevant to a case about a particular individual regardless of how much additional 'direct' evidence we had. Schoeman asks us to suppose that we know that three out of four people audited by IRS end up having to pay a lot of money. If we then find out our neighbour has just been audited and suddenly and uncharacteristically falls into a spartan lifestyle, it seems reasonable for us to make a hypothesis here, he suggests. By contrast if only one in 1000 have to pay a lot of money, says Schoeman, we would look for another reason for the change in our neighbour's situation. I am happy to grant Schoeman's point – I think statistics are relevant to making such hypotheses.[23] However, Schoeman does not leave matters at the mere making of hypotheses; he is interested in the use of hypotheses to take action regarding individuals. Schoeman suggests actions based on these assessments can equally be about individual cases as about groups. This is where my concern lies. When organisational decision-makers assess individuals, they do not simply

___

[23] From my applied ethics perspective, making a hypothesis is not directly subject to ethical judgement. Making a hypothesis is simply entertaining an idea in your mind; in itself, it is ethically nugatory. One way making such hypotheses **can** be ethically important is through influencing the character of the person entertaining the hypothesis. For example, if, when you saw people of a certain ethnicity, you repeatedly entertained the hypothesis that they were dangerous and unreliable, this might make you less tolerant as a person. In Schoeman's example, however, the judgement is one-off, not repeated. Drawing on statistics to make repeated hypotheses of this sort about individuals would perhaps only occur in a limited number of cases – if, for example, it were your job to make such hypotheses, such as a police detective investigating a crime, or an insurance assessor deciding the appropriate premium for an applicant for insurance. Perhaps these jobs can carry a moral hazard for the characters of the people employed in them, but that is outside my scope for discussion at this point. In any case, I would not expect the instance of you making a hypothesis of the sort illustrated by Schoeman, to influence your character, and so it will not gain ethical standing through that concern.

The other way simply making a hypothesis can be ethically significant is if you communicate the hypothesis to other people. This is because on receipt of the information, the other people could take action in light of it; for example, they could engage in statistical discrimination. Note that making a hypothesis does not in itself count as an instance of statistical discrimination, as it does not present a discriminatory policy or action against an individual. In Schoeman's example, however, if you communicated your hypothesis about your neighbour to others, we would probably characterise your action as gossiping. Provided you were not trying to encourage others to shun your neighbour, your voicing your hypothesis seems pretty insignificant, ethically speaking. (This need not always be so. We could imagine a lawyer, for instance, presenting a case to a jury involving hypotheses around statistical evidence. Here the presentation to others of these hypotheses can be particularly ethically significant. I address this scenario in detail in chapter 4). What does this tell us? Well, if the ethical importance of drawing on statistics to make a hypothesis about individuals comes from influencing the agent's character, or influencing others (in being communicated to others), then the instance of simply making such a hypothesis will not usually be ethically important.

hypothesise, they produce an overall assessment that is used to take action regarding the individual. I will explain Schoeman's claim, and then my concerns with it.

Schoeman suggests that not only is statistical information relevant to making a hypothesis about individual cases, but the judgements and actions we make on the basis of statistics about groups can be just as much about individual cases as about groups. He presents the example of hearing a radio report that 70% of the eggs sold at a store where he just bought eggs have been found to be contaminated with salmonella (Schoeman, 1987, p184). Given this, Schoeman asks, why wouldn't he throw them out, or have them tested, or return them for a guaranteed safe batch? But in this situation, he says, "I am making an assessment about … the eggs that I own, and not just the class of eggs generally as if this did not relate to my dozen" (Schoeman, 1987, p185). So Schoeman argues that statistical judgements can just as well be about particular or individual things, not only about classes of things: "A belief that there is a 70% probability that my dozen eggs is contaminated is a belief about my eggs" (Schoeman, 1987, p185).

It is noteworthy that the particulars Schoeman is discussing – eggs – are not individual agents. We may be quite happy accepting that Schoeman can and should make such a judgement and take action to reject his eggs in this way. I suggest it is a different matter, however, if we replace Schoeman's eggs with individual agents, such as people. Suppose that I am an employer and I hear that seventy percent of rugby-playing adults suffer from health complications that make them more likely to take time off work. Suppose too that I have a rugby-playing job applicant, Jeff. Following Schoeman's example we would continue thus: why wouldn't I throw out Jeff's application? Or force Jeff to undergo a health-test before I will even consider his application? Or require Jeff to provide his own clean bill of health before I consider his job application? My guess is that our reaction to this scenario differs from that of Schoeman's 'eggs' scenario. We do not say "but *of course* we should throw out the applications from the rugby-playing candidates, or make the candidates be tested". Unlike Schoeman's considerations of his eggs, these actions do not seem uncontroversially appropriate. I suggest the reason for this is because we are dealing with people, who are subjects with individual moral worth.

Schoeman himself says that we may have reasons for not wanting to go from a generic statistical claim (such as the percentage of rugby-playing adults with health complications) to the individual claim. But he immediately adds that "[t]his … could not be conceded without recognising that even probabilistic claims are about individuals"(Schoeman, 1987, p185). I propose the following scenario, in support of Schoeman's position. Suppose that a bank lender was required to choose between two groups of 100 people to whom to offer loans, and that all the people in each group will take up a loan if offered one. The bank lender is given no way of knowing the creditworthiness of each individual in each group. However, the lender does have the information that 70% of Group A are poor credit risks, with the remaining 30% being decent credit risks; and that these percentages are reversed in Group B. The bank lender chooses to lend to Group B. This is unfortunate for Group A, and seems unfair for the 30% of Group A individuals who are not poor credit risks. However, lending to Group A would seem unfair for the 70% of Group B individuals who are decent credit risks. In light of this, the bank lender's choice of Group B, made entirely on the basis of statistical probabilities, seems appropriate and not unfair for the individuals affected given the unusual constraints in this artificial scenario.

Note, however, in this scenario, the decision-maker is assessing one group against another group. The statistical probabilities involved are realized accurately at the level of the group, but not at the level of a member of the group taken as an individual. *Ex hypothesi*, the bank lender will get a decent return on 70 loans she makes to Group B, and no decent return on 30 loans. As individuals, members of Group B will have individual characteristics and circumstances that mean they will either default on their loan payments or they will not. This point is not reflected in the statistical probability assessment applied to any individual member of the group. For any member chosen at random from Group B, all we can say is that they have a 70% probability of being a decent credit risk, which does not reflect their individuality. Recall Standard 4 from the first section of this chapter, that a decision-maker has assessed an applicant as an individual if an assessment of the applicant uses information more likely to be true of that applicant than of any other applicant. This is not the case here.

It is true that a probabilistic claim *extends over* an individual as a unit of the group; in *that* sense it is about the individual. However, the evidence for the claim is generated from the

group only, not from the individual. The individual might not even be part of the group used to generate the claim. Consider the scenario of rugby-playing Jeff, where we have the information that seventy percent of rugby-playing adults suffer health complications that make them more likely to take time off work. Rugby-playing Jeff may have had no part in producing the percentage estimate; it is the history of other rugby-playing adults that has produced this. Moreover, we have no idea of the health-protecting and health-promoting efforts undertaken by Jeff, let alone other factors such as his work ethic, and his efficiency; it is entirely possible that Jeff is less likely to take time off work, and more likely to accomplish more work in a given period, than most other applicants. Because the evidence for the assessment of Jeff is generated from the group data only, and because the assessment (as it stands) makes no attempt to consider relevant information about Jeff as an individual, I suggest it is only very tenuously about Jeff as an individual.

In conclusion, decision-makers do draw on statistics to apply probabilistic claims to individuals. However, the fact that decision-makers sometimes apply probabilistic claims to individuals does not mean we must see them as picking out an individual rather than picking out a feature of a group. A statistical probability (or generalisation made from it) does not distinguish one member of the group from any other member, and the individual may have had no part in producing the statistical evidence on which they are being judged. In assigning a probability to a particular individual, it makes sense to see the decision-maker as simply extending a feature of a group over the individual person, rather than treating the person as an individual.

3.3 Using 'naked statistical evidence' and individual assessment to predict

The concerns about the application of group probabilities to applicants could be neutralised, however, if what we consider to be individualised assessments *also* turn out to be fundamentally based on assessment of group probabilities. This is argued by Schauer (2003, pp67-69). To focus on the conceptual issue of applying generalisations to judge individuals, Schauer (2003, pp39-41) wants to insulate it from moral reactions based on our awareness of invidious historical discrimination against groups of people. He proposes an example of

councils considering banning dangerous dogs, based on statistical generalisations about dog breeds. I will argue against Schauer. Using an example that involved individual human applicants may favour my case[24]; however, lest it be thought to illegitimately harness moral reactions, I will instead address my argument to Schauer's example. If my argument has some purchase in the context of assessing dogs, we should expect that it will apply even more strongly to the case of assessing individual human applicants.

Schauer (2003) asks what a council should make of any statistical evidence that some dog breeds, such as pit bulls, are more dangerous than others. When councils have proposed considering breed statistics, people owning dogs of those breeds have protested, calling this 'canine racism' (Schauer, 2003, p56). The protestors have noted that the majority of individual dogs of that breed do not pose a risk to people's safety, and that plenty of other dogs are more dangerous than most individual dogs of this breed (ibid). Instead of banning dogs based on breed, councils could do individualised assessments, testing individual dogs for aggressiveness. Schauer considers (2003, p65) whether councils concerned about dog violence should arrange to have individual dogs tested for their aggressive behaviour, and ban all and only dogs who test as vicious, rather than banning dogs based on group risk for the breed.

Testing dogs for aggressiveness also involves assessing individual dogs against a *group*, suggests Schauer (2003, pp67-69), namely, the group of dogs that was used as the basis for developing the test. The individual test also compares the dog against generalisations: we generalise from the actions of the group of dogs that was used as the basis for developing the test. It is on the basis of aggregate evidence from past events involving these other dogs that we say that the relevant characteristic of 'failing a test' is 'probabilistically predictive' of engaging in vicious behaviour (Schauer, 2003, p68). Basically, Schauer points out that individual testing is based on generalisations from test-performance to real-life behaviour. He declares that "using membership in the class of pit bulls as a predictor of the likelihood of aggressiveness under real-world conditions is not fundamentally different from using

---

[24] Thanks to an anonymous reviewer for *Research in Ethical Issues in Organisations* who pointed this out.

clinical testing of this pit bull as a predictor of the likelihood of aggressiveness..." (Schauer, 2003, pp66-67).

Schauer (2003, pp68-69) would also note that neither of our generalisations will produce error-free predictions. Clearly 'pit bull breed' will be an inaccurate predictor - some pit bulls would be as docile as other dogs, or more so. But tests are imperfect too, and so will also inaccurately predict for some cases: some pit bulls who pass the test will still prove too aggressive. Suppose that we have a set of three characteristics – pit bull plus male plus over a year old - that has been shown to correctly predict aggression in 85% of the dogs with those characteristics. Suppose also that we have an aggression test that, if administered to all dogs, correctly predicts aggression in 85% of the dogs that fail that test. Why should an organisation use the results of a test to exclude dogs, rather than using the fact that the dog is a male pit bull over a year old? Schauer (2003, p68) concludes that "What distinguishes the individualised examination from the so-called stereotype or the so-called profile, therefore, is only the fact that the latter is obvious without closer inspection while the former is not…".

Schauer's analysis thus suggests that with regards to generalising, a council decision-maker using membership in the class "pit bull" to judge something about a particular pit bull dog does nothing fundamentally different from a council decision-maker who does a clinical test of a particular pit bull dog. However, there's something missing here. Let us imagine a particular pit bull, "Nobbo". Suppose that this is Nobbo.

"Nobbo" [Image credit: CC_BY-SA 2.5 Alesh Houdek
http://criticalmiami.com/images/914.jpg]

Suppose too that the following image, once reproduced multiple times, represents the class of pit bulls generally.

[Image credit: Compiled by User:Oknazevad, from images from Commons, https://commons.wikimedia.org/wiki/File:Pit_bull_sampler.jpg].

Finally, suppose that this next image represents a pit bull test norm:

Schauer's alternatives seem to be presented as: 'Nobbo lumped in with the class of pit bulls' versus 'Nobbo facing individual testing'. This makes it look like we are *considering Nobbo against the pit bull class*, thus:

Fig. 1. Considering Nobbo against the pit bull class

versus *considering Nobbo against a testing norm, thus*:

Fig 2. Considering Nobbo against a test norm.

As Schauer says, if both the class and the test are generalising at base, then we could say our judgement of Nobbo is based on generalisations either way. But this wrong-foots it from the start. If a council considers a ban on dangerous dogs, and uses a classifier of 'pit bull' as

indicating a higher risk of dangerousness, *Nobbo* the pit bull is not judged. Figure 1 misrepresents the situation. This is how Figure 1 should look:
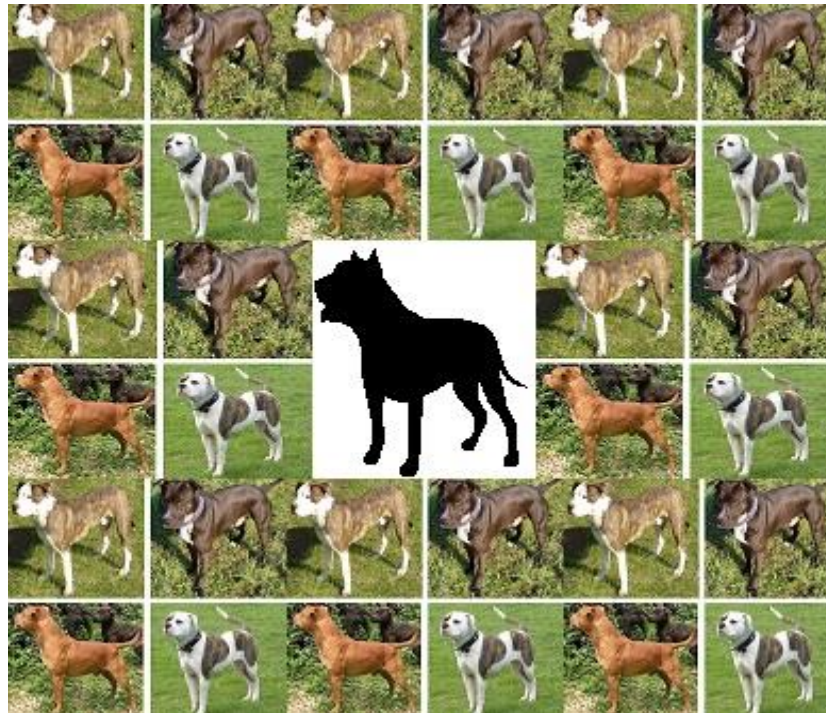

Fig. 3. Council policy banning pit bulls based on breed [silhouette image public domain: https://openclipart.org/detail/233373/pit-bull-dog-silhouette].

The crucial point is that beyond being identified as a pit bull, *Nobbo* the pit bull is given no consideration at all. Any council policy banning pit bulls based on breed risk statistics would be *applied* to individual pit bulls without being *about* this or that particular pit bull. This harks back to the point raised in the discussion of Schoeman, that class-based probabilistic claims simply *extend over* an individual. A test, by contrast, does consider the particular pit bull that is Nobbo. A test features the participation of Nobbo. It focusses on the traits *Nobbo* expresses, whether good or bad, during testing. It may also take into account *Nobbo's* past behaviour. There is thus a notable distinction between class-based and test-based assessing, as regards treating subjects as particular individuals. Judging subjects on the basis of group statistics *is* different from the test-based assessments we think of as individualised, and the failure is in the lack of consideration of the subject as an individual.

81

When a decision-maker fails to consider the people she is dealing with as individuals, she cannot help but make decisions that feature the participation of herself only (or the selves of those on behalf of whom she is tasked with making decisions). She and her principals are the only selves in the decision. She or they get to spread their selves over actions that affect others, without those others' selves being able to participate in any way. This denies any possibility of respecting the agency of the people being dealt with, and means the decision cannot but unilaterally represent the decision-maker's interests. Given this, the question of the purpose of decision-makers becomes immediately ethically significant. If the statistical claim represents the interests of the decision-makers, rather than the interests of the individual who is having the claims extended over him or her, the claims are a substantial and asymmetrical exercise of power on the part of the decision-maker.

There is already a significant asymmetry in the power organisations exert over applicants. Organisations are deciding for example, matters of liberty and autonomy (for applicants for parole, and for applicants for admission to higher education); matters of autonomy and welfare (applicants for jobs, tenancy positions and mortgages), and matters of security (applicants for insurance). Organisations can use their power to further their own interests in avoiding 'risk'. But the 'risk' that organisations associate with applicants is not confined to attributes of applicants. Organisations themselves contribute to this risk through their organisational structures, policies, and cultures. For example, Jeffrey Arthur (1994) found that the type of human resources policy practiced by firms in the steel industry impacted on the efficiency and turnover of employees. Suppose that a business organisation particularly desires having efficient, long-term employees. In this case, the business should look to how its own organisational policies and practices impact on employees, rather than just looking at the attributes of applicants. However, organisational change is difficult and bound to involve costs, at least in the short term. Organisational decision-makers will probably find it easier to locate risks in statistics attached to the attributes of applicants than to consider their own role in facilitating these risks.

In conclusion, defining risk in an organisational decision setting is in itself an exercise of power. Compared to the power wielded by organisational decision-makers, applicants are vulnerable participants in a selection process. Organisations hope to minimise risks in applicant selection through using group statistics to identify and exclude 'risky' applicants. To avoid acting hypocritically, an organisation must be willing to assess its own role in contributing to the risk it is hoping to minimise. Ideally, organisational decision-makers should be willing to explore alternatives that support lowering risks to promote more opportunity for the applicants with whom it deals. An organisation's responsibilities in this regard will depend on the resources it has at its disposal. I will explore this issue further in the next chapter.

3.4 Is the distinction between statistical and non-statistical information incoherent?

Drawing inspiration from Schauer (2003), Kasper Lippert-Rasmussen argues that what we think of as judgements of individuals are still based on statistical reasoning, in a way that collapses the distinction between 'assessment based on statistical discrimination' and 'individual assessment'. Lippert-Rasmussen is addressing the following definition of 'treating someone as an individual':

> X treats Y as an individual if, and only if, X's treatment of Y is informed by all relevant non-statistical information, e.g., information gained through X's perception of Y, other people's testimony about Y delivered to X, X's memory of what Y did in the past, available to X. (Lippert-Rasmussen, 2011, p51)

He suggests this definition is unsatisfactory because the distinction between statistical and non-statistical information can reasonably either be denied or be seen to be completely superficial because non-statistical information is based so thoroughly on statistical information.

Lippert-Rasmussen (2011, p49) asks us to compare the reasoning of two employers, A and B:

(A) ''I seem to observe that this applicant is a 25 year old woman. Hence, it is almost certain that she is a 25 year old woman. 30 out of 100 twenty-five year old, employed women apply for parental leave within 5 years. If I hire her, the probability that she will apply for parental leave is almost 3–10.''

(B) ''I seem to observe that this applicant is telling me that she intends to have a child within the next 5 years and apply for parental leave. Hence, it is almost certain that this is indeed what she is telling me. 30 out of 100 applicants who say they intend to have a child and apply for parental leave within the next 5 years apply for parental leave within 5 years. If I hire her, the probability that she will apply for parental leave is almost 3–10.''

Note the difference we are supposed to remark here: that employer B is taking account of the personal testimony of the applicant. In terms of my analysis, Lippert-Rasmussen's example differs from Schauer's because taking a statement from an individual means the assessment process does involve the participation of the individual.

Lippert-Rasmussen (2011, p51) suggests that at first glance, employer B seems to be treating the applicant as an individual whereas employer A does not[25]. But he submits that, on reflection, it is hard to see any significant differences between the two types of reasoning. He presents Schauer's point that "… seemingly direct observation involves a process of inference and generalization …. even the processes that initially appear to us to be 'direct,' 'actual,' or individualized turn out to rely far more on generalizations from past experience than is often appreciated …. all evidence is in the final analysis probabilistic…" (Schauer, 2003, cited in Lippert-Rasmussen, 2011, p51). Lippert-Rasmussen implies that statistical generalising goes 'all the way down' in our decision-making. He concludes (2011, p51) that any distinction we draw between treating someone on the basis of individualised

---

[25] In a footnote on the section on Schoeman earlier in this chapter, I noted how making a hypothesis differs from engaging in statistical discrimination. Neither of Lippert-Rasmussen's examples show an employer engaged in statistical discrimination; as it stands, they show the two employers making a hypothesis. For an example to show statistical discrimination, the employer would need to take some action toward the applicant on the basis of the conclusion of the statistical reasoning. In what follows, I will presume the employers do this, in order to discuss the examples in terms of statistical discrimination.

evidence and treating them on the basis of statistical generalisation is either superficial or ultimately incoherent.

Lippert-Rasmussen is right that that both employer A and employer B engage in statistical discrimination against the applicant. However, the 'Employer B' example does not relevantly reflect a decision-maker making an individualised assessment in this situation. So the fact that employers who engage in the sort of reasoning that A and B do are both, at base, engaging in statistical discrimination does not take Lippert-Rasmussen's argument very far. Employer B does a speech-to-action probability calculation, on the basis of statistics that have nothing to do with the individual. Lippert-Rasmussen suggests that initially, it looks like Employer B treats the applicant as an individual: he would have us focus on the point that Employer B makes note of personal testimony from the applicant. However, I would draw attention to the point that the information Employer B uses in assessing the applicant is *just the group statistic*. Moreover, with a statistic indicating only a 3 in 10 chance, Employer B would need to judge that it is more likely false than true that the applicant will carry through on her intention. If Employer B *did* factor in this statistic in making a decision, Employer B would use it to *dismiss* the applicant's stated intention. I suggest this looks much less like the decision-maker treating the applicant as an individual.

Lippert-Rasmussen's point with his 'employers' example was intended to apply to interpretations of speech-to-action probability across the board. That is why he suggests, on the strength of this example, that the distinction we draw between treating someone on the basis of individualised evidence and treating them on the basis of statistical generalisation breaks down. Lippert-Rasmussen suggests that the decision-maker's judgements using 'individualised' statements of intention simply reflect statistical discrimination operating at another level - some kind of background statistical discrimination that we should see as underlying all interpretations of speech about intentions.[26] I am suggesting that the concept

---

[26] In his 2014 chapter on statistical discrimination, Lippert-Rasmussen refers to the type of statistical discrimination from Employer A as being 'foregrounded', whereas the speech-to-action probability of Employer B is 'backgrounded'. He ends up suggesting that if we admit 'backgrounded' statistical discrimination – that is, if we admit speech-to-action probability assessment as statistical discrimination - then we face a 'limitation' problem, as too many acts will count as discrimination.

of treating someone as an individual by using 'individualised' evidence is not compatible with simply dismissing the individualised evidence based on group statistics.

I think Lippert-Rasmussen would agree with me that ignoring available individualised evidence and using just statistical generalisations is not treating someone as an individual. He characterises himself as a 'revisionist' on the issue of treating people as individuals: "Revisionists do not contrast treating people as individuals with treating them on the basis of generalisations. Rather, revisionists hold that failing to treat someone as an individual involves a particular way of treating people employing certain general information while ignoring other pieces of information." (Lippert-Rasmussen, 2011, p53) Accordingly, he suggests the following definition of what it is to treat someone as an individual: "X treats Y as an individual if, and only if, X's treatment of Y is informed by all relevant information, statistical or non-statistical, reasonably available to X." (Lippert-Rasmussen, 2011, p54) This definition allows that the consideration of information gained through statistically-backed generalisations will not necessarily result in decisions that fail to treat people as individuals[27].

Lippert-Rasmussen's definition of treating a person as an individual may be problematic. If X has all relevant non-statistical information reasonably available to X, then if X adds statistical information to this, this would seem only to decrease the possibility of X treating Y as an individual. For example, suppose that X is an employer and Y is a job applicant. X has testimony from two previous employers that Y has a good work attendance record. Now suppose that X finds credible research that only 30% of group 1, of which Y is a member, has good work attendance records. Alternatively, suppose that X finds credible research that fully 85% of group 1, of which Y is a member, has good work attendance records. Using either of these statistics to inform X's decision seems only to detract from treating Y as an individual. Presumably, Lippert-Rasmussen could respond that, in the case of sound non-statistical evidence being available, the statistical evidence should be considered no longer relevant. I would be happy with that response.

---

[27] Lippert-Rasmussen's definition does also cover a situation where X has only statistical information relating to Y (where no non-statistical information is available). However, I think it makes more sense to doubt that X can treat Y as an individual in such a situation.

This response aligns with Benjamin Eidelson's account of treating a person as an individual. Eidelson (2013, p375) suggests the following as a formal account of treating people as individuals:

> In forming judgments about Y, X treats Y as an individual if and only if:
>
> (*Character Condition*) X gives reasonable weight to evidence of the ways Y has exercised her autonomy in giving shape to her life, where this evidence is reasonably available and relevant to the determination at hand; and
>
> (*Agency Condition*) if X's judgments concern Y's choices, these judgments are not made in a way that disparages Y's capacity to make those choices as an autonomous agent.

Eidelson sees a strength of his account in its ability to distinguish between the two scenarios from Lippert-Rasmussen, evaluating the first as unacceptable, while the second is acceptable (Eideslon, 2013, p380). In the first scenario, the employer forecasts what the applicant is likely to do based simply on her age. In so doing, the employer pays no heed to evidence of the ways the applicant has exercised her autonomy in shaping her life, and disparages the applicant's capacity to make a choice as an autonomous agent. In the second scenario, by contrast, where the employer takes a statement from the applicant and runs a statistical forecast from it, the employer pays attention to evidence from the applicant.

Because Eidelson's (2013, p375) definition requires the decision-maker to give a reasonable weight to evidence of Y's past chosen actions and future choices, Eidelson's account can recommend the exclusion of purely statistical evidence where there is sound individual evidence available. Conversely, it can appropriately allow exclusion of individual evidence (and space for the use of high statistical probabilities) when the individual evidence is of poor quality. For example, suppose that X has heard that both of Y's previous employers lack integrity and have a reputation for positively inflating references for poor employees to help move the poor employees on to other employers. In this case, X can reasonably give much less weight to Y's testimonials.

Eidelson's position on what it is to treat someone as an individual may not deal with concerns about the treatment of rugby-playing Jeff, however. Recall the employer rejecting

the application of rugby-playing Jeff due to the statistic that 70% of rugby players develop health problems that make them more likely to take time off work. Jeff must have exercised his autonomy in deciding to play rugby – this is obviously a chosen activity. If a decision-maker rejects Jeff out of hand due to the rugby statistic, could the decision-maker be said to have given reasonable weight to evidence of a way Jeff exercised his autonomy in giving shape to his life? Whether this is reasonable is contestable and may be hard to answer– it may seem reasonable from the point of view of the decision-maker, but not from the point of view of the applicant; and society may not hold a united perspective on this either. To try to give an answer to this, I think we need to consider further the purposes and interests of both parties, and society generally. I will undertake this task in chapter six of the thesis.

Eidelson's position does not recognise that *non-autonomous* features of persons can be part of their individuality. I am 169 centimeters tall; the fact "Vanessa Scholes is 169 centimeters tall" is a fact about the individual I am. If I were sick with a virus, the fact "Vanessa Scholes is sick with a virus" would also be a fact about me as an individual. So individuality can encompass more facts than just what someone autonomously chooses to do. However, Eidelson wants to explicate the morally important sense in which people are individuals, which he thinks resides in a person's autonomy to author her own life (Eidelson, 2013, p356), so naturally his position focusses on this. I agree with Eidelson that this is morally important. I also think there is more that is valuable about treating people as individuals, and this will be my focus in the final section of this chapter.

3.5 What is valuable about treating people as individuals?

The foregoing discussion implies that there is something valuable about trying to treat people as individuals, that risk-based statistical discrimination does not allow for – something that could perhaps make it morally preferable to try to treat people as individuals. I think this is true, because I think that this is tied to valuing people as *unique autonomous agents*. I suggest that trying to treat people as individuals is important in part because it allows us to prioritise recognising a person as a locus of unique moral agency, which I think has moral worth. The claim that there is moral worth to recognising an

individual person as *unique* expresses the idea that the non-substitutability of each person is a source of value. This is partially captured by the idea of a person being an individual 'in his or her own right'. The claim that there is moral worth to recognising a person as a locus of moral agency expresses the idea that the autonomous moral agency of a person is a source of value. This is captured by the Kantian idea of valuing people being able to work toward achieving their own ends (agency) in part through setting and following principles for themselves (autonomy).

These valuings of people may conflict with assessing applicants using risk-based statistical discrimination to disadvantage or exclude applicants. Assessing applicants in this way may ignore their individuality. First, by categorising applicants as a member of just a couple of groups and disadvantaging or excluding them on this basis, a decision-maker ignores each applicant's uniqueness. Second, if applicants are categorised on the basis of group features that do not involve the applicants' (past or present) chosen actions, this ignores the applicants' autonomy. Third, if applicants are categorised on the basis of group features with no obvious causal connection to the organisational good, this stymies the applicants' agency. For example, suppose that there are many applicants for IT jobs in Jordan, and applicants take pains to study appropriate courses, work on IT projects in their spare time, build their portfolios of skills – things that have an obvious causal connection to securing and doing well at an IT job. However, computer algorithms detect a reliable correlation between opera attendance and better work performance in the Jordanian IT industry[28]. If employers used the factor of 'opera attendance' to screen applicants, this would work to forestall applicants' agency because it is not clear how opera attendance improves IT capabilities. If the job is part of the applicants' own ends, then the use of features statistically correlated with job performance, but bearing no obvious causal connection, blocks the pathways through which applicants might realise this end, and so stymies their agency.

---

[28] This example is fictitious; however, this is the sort of thing that computer algorithms, and in particular machine learning algorithms, can do, and I will be discussing developments of this sort in later chapters (particularly in chapter 6 and in my conclusion).

Lippert-Rasmussen (2007, p395) disputes the idea that there is something morally preferable about treating people as individuals:

> The attractiveness of the view that people ought to be treated on the basis of their individual properties is due to the belief that if people are being treated on the basis of properties they are reasonably believed to have, it will be more likely that they will receive the treatment that they ought to receive given the individual properties they in fact have.

Lippert-Rasmussen proceeds to note that this is not necessarily the case – that use of statistical discrimination in some instances could mean people are more likely to receive treatment matching the individual properties they have. This could happen, for example, where decision-makers tend to draw on inaccurate stereotypes in decision-making. In this case, substituting statistical grounds for decisions would mean more of the inaccurately-stereotyped group would get treatment matching the properties they have. However, I think Lippert-Rasmussen offers an impoverished description of the attractiveness of the view of what is morally preferable about treating people as individuals. I will draw on work from David Edmonds (2006) and Benjamin Eidelson (2013) to help explain why.

Edmonds (2006) argues that our judgements of the moral acceptability of discrimination on statistical correlations hinge on an interplay of two factors. One is our assessment of causation (or internal conceptual correlation) and the other is responsibility or desert. We think it morally more appropriate to categorise people on the basis of what they have done (or sometimes failed to do) and where the doing of a thing is "directly, causally linked in the appropriate fashion" (Edmonds, 2006, p33). This suggests that the unattractiveness of the practice of statistical discrimination is likely to vary in accordance with whether the group feature involves an action taken by members; and whether we think there is a causal connection between the action and the feature. For example, when assessing a person for auto insurance, discriminating against an applicant for their poor credit history[29] should be more attractive than discriminating against an applicant for being a young man, because the applicant is not responsible for the latter feature. Yet discriminating against an applicant for

---

[29] Federal Trade Commission (July 24, 2007). FTC releases report on effects of credit-based insurance scores. Retrieved 26 July 2016 from https://www.ftc.gov/news-events/press-releases/2007/07/ftc-releases-report-effects-credit-based-insurance-scores

their poor driving history should be still more attractive than the credit-based discrimination; since poor driving history is both based on effort and directly causally linked to the issue under judgement (auto insurance). So the attractiveness of individual treatment is not just if it might better pick out properties people have (such as their history) but also it is more likely to include a focus on the causal connection between the properties and the treatment of the individual.

Moreover, the attractiveness is not, or not just, the consequences suggested by Lippert-Rasmussen. The attractiveness of valuing treating people as individuals includes how this lets people interact. It means people can be treated in a way that recognises their history, that privileges their past and current efforts and demonstrated abilities and capabilities over other information. Valuing people as individuals also means the decision-maker can see a person's future actions as open to being decided by that person; that is, the decision-maker can see the person's future as open to them choosing actions that may or may not be congruent with their past[30]. In short, as noted by Eidelson (2013), it takes account of the person's past and future autonomy and agency. It also means the decision-maker, where relevant, can take some account of the person's self-perception. The decision-maker can thus show a degree of epistemic humility and awareness of the influence of socio-cultural and psychological assumptions on their judgements (e.g., their judgements of the capabilities of a disabled person). This means it allows the conditions necessary to treat a person as a locus of unique moral agency.

Conclusion

I have given moral reasons to support a preference for organisational decision-makers to treat people as individuals. These are grounded in respect for a person as an unique moral agent: a person whose individual autonomy and agency are a source of value. This respect can be implemented through assessing an applicant with appropriate use of individualised evidence. This will include focussing on an applicant's past and present actions that stand in a causal relation to their capabilities regarding the organisational good. This respect can also be expressed through considering the influence of the organisational environment on risk

---

[30] Many of these points are captured in Eidelson's (2013) autonomy account of treating people as individuals.

factors regarding applicants. Note that this does not rule out some use of statistical evidence. However, I argue that an organisational decision-maker who uses a few features statistically correlated with higher risk to dismiss or downgrade an applicant is not treating the applicant appropriately as an individual, and that this is morally problematic. I see this as a reason to morally disprefer excluding or disadvantaging people purely on the basis of statistics for a group of which they are a member.

However, autonomy and agency tend to be more important in some decision-contexts than others. They are very important in life-and-death medical decisions, and not very important when a tour guide is separating customers into groups for a half-hour tour. As noted in chapter 2, not all social contexts are important to our individual identity. Perhaps when what is at stake is relatively unimportant to our individual identity, we will not see statistical discrimination as threatening our autonomy and agency. Moreover, respect for an applicant's autonomy and agency is only one relevant ethical consideration when looking at organisational assessment of applicants. Clearly, the consequences of using or not using pure statistical evidence, for an organisation, the applicant and other people, are ethically significant. In chapter four I take up some of these issues. I look at how the purposes of organisations and the resources organisations employ can affect the moral status of statistical discrimination.

Chapter Three References

Arthur, Jeffrey B (1994) Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal, 37* (2): 670-687.

Baumle, Amanda K. and Fossett, Mark (2005) Statistical Discrimination in Employment: Its Practice, Conceptualization, and Implications for Public Policy. *The American Behavioral Scientist, 49*: 1250-1274.

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York: Routledge.

Eidelson, Benjamin (2013) Treating people as individuals. In Deborah Hellman and Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp354-395) Oxford University Press: Oxford.

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, MA: Harvard University Press.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *The Journal of Political Philosophy*, 15: 385-403.

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15: 47-59.

Lippert-Rasmussen, Kasper (2014) Statistical discrimination. Ch. 3 In *Born Free and Equal: A philosophical inquiry into the nature of discrimination*. Oxford Scholarship Online DOI:10.1093/acprof:oso/9780199796113.001.0001

Maitzen, Stephen (1991) The Ethics of Statistical Discrimination. *Social Theory and Practice*, 17 (1): 23-45.

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Schoeman, Ferdinand (1987) Statistical vs direct evidence. *Nous*, 21 (2): 179-198.

Scholes, Vanessa (2016) Risky business – the ethics of judging individuals based on group statistics. In Marco Grix and Tim Dare (Eds.), *Contemporary Issues in Applied and Professional Ethics* (pp. 169–188). Emerald. doi:10.1108/s1529-209620160000015010

Chapter four   The importance of context: purposes and resources

Introduction

In this chapter I explore the decision-making context for the practice of statistical discrimination, and how it affects the ethical permissibility of an organisation's treatment of individuals. I argue first that the ethical permissibility of the practice of statistical discrimination varies depending on the purpose of the organisation in society. I illustrate this by contrasting statistical discrimination in two different settings: courts of justice, and insurance. I argue that an organisation's purpose bears on the ethics of statistical discrimination in terms of the fit between the ends of the organisation and the assessment method. In the case of a court, employing statistical discrimination seems contrary to the ends of the organisation, whereas in the case of insurance companies, it seems indispensable to their proper functioning.

Second, I argue that the risk-based statistical discrimination employed by insurers is less discriminatory than such discrimination in other commercial areas. Comparing insurance with an example of a tenancy agent employing risk-based statistical discrimination, I argue that the statistical discrimination by insurers is less 'discriminatory' because it can treat applicants more as individuals. It is also less morally problematic as discrimination, for two reasons. The minor reason is that it can produce beneficial information for all applicants as part of the process. The major reason it is less morally problematic is because the transparency of the organisation's purpose means applicants know that they will be treated on the basis of statistical generalisations, so they can give informed consent to this treatment.

In the third, smallest part of this chapter, I argue that the ethical permissibility of the practice of statistical discrimination will depend on the resources the agent has available to them to influence the informational context, especially the resources they have to spread risk. Risk spreading can alter the need to rely on group-risk statistics to manage risk, changing the incentive for statistical discrimination. I discuss "all-or-nothing" decisions, a general decision-making context that Frederick Schauer argues tends to support the use of

statistical discrimination. I will argue that whether a context is or becomes an "all-or-nothing" affair can depend on the resources available to the decision-maker to spread risk, and that applicant selection decisions may offer the opportunity for this. I conclude that the purpose of organisations and the resources they have to influence their informational contexts are key contextual factors influencing the ethics of statistical discrimination.

4.1 The purpose of Courts: 'naked statistical evidence' and justice

The 'Blue Bus Problem' is a hypothetical case used in teaching law that is based on a true case (Schauer, 2003, p81ff).[31] The situation involves a plaintiff who has been hit by a vehicle on a night so dark that the only characteristics of the vehicle she can identify are that it was a bus and it was blue, and a defendant, the Blue Bus company, who owns 80% of the blue busses in the town; these facts are not in dispute (Schauer, 2003, pp81-82). The plaintiff takes a civil case against the Blue Bus company. Unlike a *criminal* case where guilt must be established 'beyond reasonable doubt', in a *civil* case, all the plaintiff needs is for the 'preponderance of evidence' to favour her case against the defendant. 'Preponderance of evidence' is usually understood, says Schauer (2003, p81) as the 'balance of probabilities' or as "just over a 50 percent likelihood that the proposition asserted is true." On the 'naked statistical evidence' there is an 80% probability that the blue bus that hit the plaintiff was a Blue Bus company bus, which seems to well exceed the 'preponderance of evidence' standard required. This would suggest the court ought to find in favour of the plaintiff. Yet, Schauer notes (2003, p81), the courts routinely dismiss cases like this where the evidence is based entirely on statistical probability.

Schauer offers an argument in favour of the use of (some) 'naked statistical evidence' in courts in view of problems with a key 'individualised' alternative to this, namely use of witness testimony. Due to known problems with perception and bias, witness testimony is notoriously unreliable (Schauer, 2003, pp94-96), so use of witness testimony about X may be less likely to produce a reliable outcome in courts than the use of statistics about the background probability concerning X. Think back to our earlier scenario of Jane Combs, found comatose due to food poisoning from eating fast-food. Suppose we have two cases.

---

[31] The Blue Bus problem was originally formulated by Professor Laurence Tribe (Guerra-Pujol, 2014, pp46-47).

In the first, Jane lives in Old Buckshire which has a varied range of fast-food outlets. We have an eyewitness report, recorded two days after the event, of Jane entering a Happy Harry's Organics-to-go cafe on the morning she went into the coma. Suppose, however, that problems with perception and memory mean eyewitness testimony has only a 75% probability of being accurate. In the second case, Jane lives in New Buckshire and we have no eyewitness reports of her movements on the morning in question, but we do have the statistic that 80% of all fast-food outlets in New Buckshire are Happy Harry's Organics-to-go cafes. Comparatively, it seems that making a judgement using the high percentage in the New Buckshire case is a better bet than using the eyewitness testimony in the Old Buckshire case. Let us call this the 'better bet' argument.

In isolation, the 'better bet' argument looks impressive. However, if we measure the argument against the purpose of a court, some problems appear. We have a justice system to serve justice through enforcing the law. A court, as part of the justice system, has the purpose of serving justice in relation to the particular cases brought before it.[32] Courts need to be independent, fair and efficient.[33] Tom Tyler (2014, p1095) identifies two distinct objectives of courts: "establishing the truth and punishing justly".[34] Establishing the truth requires having a method that does best at framing the particular situation being complained about. 'Punishing justly' is a course of action that appropriately disciplines and deters an appropriate party, whether defendant or plaintiff[35]; hopefully, it will also send an

---

[32] Canadian Superior Courts Judges Association "The justice system is the mechanism that upholds the rule of law. Our courts provide a forum to resolve disputes and to test and enforce laws in a fair and rational manner." Retrieved from http://www.cscja-acjcs.ca/role_of_courts-en.asp?l=4

[33] Courts of New Zealand "Independent, fair and efficient courts are an important cornerstone in our democracy." Retrieved from https://www.courtsofnz.govt.nz/about/system/role/overview

[34] Cf Ronald Allen's claim that "Evidence law should facilitate the accurate, efficient, and fair finding of facts pertinent to legal disputes." (Allen, 2015, p77). Serving justice might include other aims such as setting an appropriate precedent, but presumably this is not the primary aim.

[35] Wesley Cragg (1992) suggests the use of the authority of legal systems "is morally justified … when it reduces recourse to the morally justified use of force in dispute settlement." ….] [….] "The function of a sentence, this account suggests, is the resolution of disputes to which criminal offences give rise in ways designed to sustain confidence in the capacity of the law to fulfil its legitimate functions on the part of victims of crime and the public at large. To do this effectively, the sentencing process must demonstrate commitment on the part of authorities to law enforcement, while seeking to persuade and enable offenders to live within the law." (Cragg, 1992, p7) Hallevy (2013, p16) identifies the four general purposes of punishment as Retribution, Deterrence, Rehabilitation and Incapacitation.

appropriate message of discipline and deterrence to society.[36] In a civil case, discipline and deterrence could be imposed either on the defendant, if found guilty, in the form of liability to pay damages and court costs; or on the plaintiff, if the defendant is found not guilty, in the form of liability to pay court costs. Under this conception of serving justice, I will argue that the use of 'naked statistical evidence' of the sort presented in the Blue Bus problem case or Jane Combs's case is inimical to serving justice, as it fails to focus on the particular case and fails to be an appropriate basis for discipline, which is unfair to the defendant. It also fails as an appropriate basis for deterrence.

Suppose that Blue Bus company owns all the blue busses in the city, and 75% of the bus stock overall. Suppose further that the remaining 25% of busses are owned by the Green Bus company, whose stock comprises green busses. If we allow the use of naked statistical evidence, then in any particular instance where a bus impacts someone who cannot make out the colour (say, at night), the balance of probabilities says that we should judge it to be a Blue Bus bus. However, over a period of busses hitting people under such conditions, and absent any other information, we could expect that 25% of the time it is a Green company bus that hits someone. If the method of using naked statistical evidence is recommended for judging cases in courts, then it seems we should specify that 25% of the time in a case of this sort – in other words, in every fourth court case of this sort - the judge or jury should conclude it was a Green company bus.

The very fact we could specify *in advance* that in every fourth case of this sort, Green Bus company should be held to be the guilty party, shows the method has no connection to what happened in any particular case. David Enoch and Talia Fisher (2015) assess naked statistical evidence using the concept of Sensitivity – "the requirement that a belief be counterfactually sensitive to the truth" (Enoch & Fisher, 2015, p559). Naked statistical evidence fares poorly because it precludes the counterfactual – if it *were* the case that a

---

[36] "Looked at as a whole, the penal system embodies a range of aims, from retribution, punishment and deterrence to reform and rehabilitation, protection and prevention. [….] most penal measures probably represent one aim more than others and most aims probably characterise some penal measures more than others. Thus, fines and imprisonment are associated with the aims of punishment and deterrence more than any others; community service is associated with reparation to society and making the offender aware of the effects of crime on the victim…" (Lord Low of Dalston, 16 May, 2007 4.45pm)

Blue bus did *not* hit anyone, statistical evidence will still compel us to judge that it *did*. This is unfair for the defendant – they are basically judged guilty in advance of consideration of their particular case. Even if naked statistical evidence does better at predicting the truth about what happens over a series of colour-blind bus impacts, this is not what it is to consider the particular case at hand.

It is noteworthy that the particularistic judgement required in court cases involves a *backward-looking* aspect, focussing on whether someone did something in the past. Could this be part of the problem? Perhaps naked statistical evidence is simply unsuitable for backward-looking judgement. This point is mentioned in passing by Lippert-Rasmussen (2007, p391, footnote 14).[37] Using statistics derived from the relative frequencies of a class of events may be better suited to decisions where agents are intent on making predictions about what is likely to occur in the future, rather than decisions where agents are intent on establishing what occurred in the past.

Naked statistical evidence fares no better if we consider the aims of discipline and deterrence. Individuals could be rewarded or punished on the basis of statistics that do not reflect the individual's actions. This is patently unfair, and no fit basis for discipline. This also gives the individual party no reason to change their particular actions, and so is no fit basis for deterrence. If a practice of using naked statistical evidence were adopted, instead of deterring a party found to be responsible, it may provide perverse incentives for parties. Consider again the Blue Bus scenario. If there are frequent court cases with courts awarding bus victims large sums of money against Blue Bus company on the basis of naked statistical evidence, this could serve to incentivise Blue Bus to reduce their bus ownership to 50% so they would be immune from prosecution on the balance of probabilities. If judgements based on naked statistical evidence are not a fit basis for imposing discipline and do not deter, such judgements do not seem to fit the requirements for courts to impose fair punishment.

---

[37] "The use of statistics in legal settings typically differs from its use in the sort of cases I focus on here. First, it serves to establish what happened specifically at a certain moment of time and not to determine the relative frequencies of a class of events." (Lippert-Rasmussen, 2007, p391, footnote 14). Schauer does not make a point of distinguishing between backward and forward-looking judgements, moving from examples in one context to examples in the other without pause (2003, p96).

The 'better bet' argument is not the only argument that could be put forward in favour of the use of naked statistical evidence in the legal setting, however – **efficiency**, too, could be noted. It is quite plausible that the more we want to insist on individualised information the more we will need to sacrifice some efficiency. Consider, for example, the time and trouble of gathering evidence and rounding up appropriate witnesses, versus the time and trouble of producing a statistic such as that Blue Bus company owns 80% of the city's busses. The use of naked statistical evidence might speed up court decision processes, allowing a greater number of court cases to be processed in a given time. Taking efficiency into account may sound like a departure from the idea of courts serving fair punishment, if the suggestion is that we focus on what is quick rather than what is fair. However, in the real world courts may have backlogs of cases to be decided. If 'justice delayed is justice denied', then efficiency could be a point in favour of the fairness of the use of naked statistical evidence.

Moreover, perhaps we should not worry about not getting the full picture of a particular case, since serving fair punishment in a particular case may not achieve justice in a wider sense. If a poorer person steals from a richer person in a particular instance, and the richer person gets a court to serve fair punishment on the poorer person resulting in the near-destitution of the poorer person, that might not be justice in a wider sense. While the wider conception of justice must undergird the acceptability of courts of justice and their practices in some sense, this does not mean the courts need to take account of the individual circumstances of each party before it. Schauer notes (p103) that use of individualised rather than statistical evidence to assess parties will not give us the full individual picture of any particular case, and it will leave out information that might be relevant. For example, judges do not allow consideration of the plaintiff's or defendant's financial standing, insurance coverage, or past negligent acts - regardless of their relevance to serving justice, more broadly speaking (Schauer, 2003, p105).

If we accept Schauer's points that courts need not consider fully individualised pictures of the parties before them, and that greater individualisation is less efficient, this seems to press us for a justification for insisting on the 'given' level of individualisation of information

in assessing parties in legal settings. However, there is a reason courts need not consider fully individualised pictures of the parties before them in order to serve a broader conception of justice. This is because courts are just one part of the set of institutions that serve this broader conception. Courts are one state institution set up to promote justice; they are not the sole arbiter or dispenser of justice in society. There are other state institutions (such as social welfare and public healthcare systems) and also non-state organisations (NGOs, churches), that contribute to achieving a broader conception of justice in society. This means we need to hark back to the purpose of courts in serving justice in society. What is the purpose of a court in relation to justice in society, and is it compatible with more efficient, non-individualised evidence?

I have suggested that a court's role (at least in part) should be to serve fair punishment in particular cases. Whatever else we might want from a state's justice system, I think we do want an institution that does this. To see why, we can take Schauer's points on efficiency and non-individualised information and push them further. Perhaps there is a more efficient way of aiming for overall justice: we could simply dispense with the court system altogether. In its place, we could use our best victim statistics to decide the likely numbers of victims across a series of broad social groupings (such as age, gender and ethnicity) in a given year; and likely number of perpetrators, again using broad groupings (such as age, prior convictions). We could then fine or otherwise sentence a statistically accurate subset of the 'perpetrators' groupings, and send letters of victim support or monetary compensation to a statistically accurate subset of the 'victims' groupings at the end of each year. This would certainly be more efficient than running a court system. However, regardless of efficiencies, I suggest this departs so far from our idea of fair punishment that it cannot be part of a state's justice system.

I have argued that application of naked statistical evidence to determine outcomes of court cases is contrary to an idea of justice that comprises courts serving fair punishment. Naked statistical evidence's failure to focus on the particular case is unfair to the defendant. In these judgements, I suggest, *naked* statistical evidence is indecent. This is not to deny a role for any statistical evidence - it could be decently admitted, for example, as a factor bearing on the plausibility of particularistic evidence, such as noting witness reliability issues when

witness testimony is presented. I do, however, deny its acceptability as a replacement for particularistic evidence. The purpose of a court is to serve fair punishment in particular cases. Naked statistical evidence, evidence that is derived from the relative frequencies of a class of events, does not suit a backward-looking particularistic judgement, such as we see in court decisions. When an organisation has such a purpose, statistical evidence must be decent, not naked.

I have argued against the fairness of statistical discrimination based on naked statistical evidence in cases before courts, due to the purpose of courts of justice. My overall aim in this chapter is to argue that the ethical permissibility of risk-based statistical discrimination varies according to the purpose of the organisation in society. To support this claim, I need offer some arguments in favour of the fairness of statistical discrimination by an organisation with a different purpose. As noted earlier, statistical discrimination may be better suited to decisions predicting what is likely to occur in the future, rather than decisions where agents are establishing what occurred in the past. In looking forward to assess the future, a decision-maker is not necessarily aiming for truth, discipline or deterrence; instead, he or she is aiming for a useful forecast. As luck would have it, there is an industry based on using naked statistical evidence to provide forecasts about individuals: the private insurance industry. I will present and discuss arguments for the fairness of statistical discrimination by organizational decision-makers in private insurance, drawing on the work of Joseph Heath. I will also argue that this statistical discrimination in insurance may be less 'discriminatory', and less morally problematic as discrimination, than in other commercial areas.

4.2 The purpose of insurance : actuarial assessment and discrimination

The insurance industry works on the principle of discriminating between people on the basis of statistical correlations for groups to which they belong, using this to forecast risk. The information applicants provide is mapped against group statistics for particular characteristics to form an overall risk profile – an actuarial classification - used to determine

whether applicants will be offered an insurance policy, and at what price.38 For example, post-natal depression is a risk factor for health insurance, because women who have had post-natal depression, as a group, are more likely to have subsequent depression and increased healthcare claims in future. So insurers will want to know about this factor in order to offer these women worse terms for their insurance than they would have been offered had they not had post-natal depression. Similarly, when assessing a person for automobile insurance, underwriters may look at a person's credit history, because drivers with poor credit histories, as a group, make more automobile claims against insurers.39 So 'bad credit history' is a statistical indicator for worse risk of making an automobile insurance claim.

Risk classification is the most basic principle of insurance (Stone, 2002, p73). Insurers make statistical discriminations in order to charge people different insurance premiums, based on their risk. Insurance applicants from groups which are statistically more likely to require a payout, or a higher payout, are charged higher premiums. Conversely, insurance applicants from lower-risk groups are charged lower premiums. People can make choices that reduce or increase their overall risk, which can affect their premiums. For example, if you have house insurance and you choose to install a house alarm, this can reduce your premium. Dangerous recreational activities such as skydiving or mountain climbing attract increased premiums; people can reduce their risk, and so reduce the amount of insurance they pay, by choosing less dangerous activities. Avraham, Logue and Schwarcz (2013, pp2-3) suggest that accurate risk classification and incentivising risk reduction are the two main reasons we let insurers discriminate.

Economists believe that, for reasons of economic efficiency, individuals should be charged premiums on the basis of the risk the individuals bring to the insurance arrangement. From an economic viewpoint, if an individual can be charged a premium that reflects the

---

[38] Some parts of this paragraph are taken from Scholes (2014).
[39] Federal Trade Commission (July 24, 2007). FTC releases report on effects of credit-based insurance scores. Retrieved 26 July 2016 from https://www.ftc.gov/news-events/press-releases/2007/07/ftc-releases-report-effects-credit-based-insurance-scores

expected cost of compensating the individual for the loss, multiplied by the probability of the loss occurring[40], this would constitute an ideal of 'actuarial fairness' (Heath, 2014, p345). Legislators have also agreed that it is fair (or fair enough) to set insurance premiums on the basis of group risk. Amanda Baumle and Mark Fossett (2005) looked at insurance in some North American states. They suggest that legal rulings and public policies on insurance have also conceived of 'fairness' as 'fairness at the group level'. Suppose that the group 'male drivers under 25 years old' is involved in more car crashes than any other group; as long as the insurance premiums accurately reflect the average crash risk for that group, those premiums have been considered fair (Baumle & Fossett, 2005, pp.1268–1269). By contrast, Heath (2014, pp345-6) suggests that others, including philosophers and civil rights law theorists, often seem to reject the notion of actuarial fairness and hold to an ideal of 'community rating' where individuals' risk profiles are ignored and everyone pays the same price for the same policy[41].

---

[40] For example, suppose that Warren is a 21 year old male with a 10 year old car of a type that is often used for 'boy racer' activities, and Ava is a 51 year old woman with a 10 year old Lexus. Drawing from the statistics associated with these features (along with others – I am simplifying here), insurers work out a premium.

  - If the costs associated with Warren having a car crash were $20,000, and if the probability of Warren having a crash was 5%, then his premium should include a cost of $1000 for his car-crash propensity.

  - If the costs associated with Ava having a car crash were $20,000, and if the probability of Ava having a crash was 0.5%, then her premium should include a cost of $100 for her car crash propensity.

[41] Community rating has some drawbacks. It does not incentivise insurance customers to minimise their risks, as there is no insurance benefit to customers from doing this. Why spend time and money putting in a house alarm and maintaining it when it will not reduce your insurance premium, and your insurance means you will be paid for anything stolen? Another concern with community rating is 'adverse selection', which is when higher proportions of higher risk groups select the insurance. Adverse selection is more likely with community rating, as the insurance is a much better deal for people in higher risk groups. The lack of incentives to reduce risk and the issue of adverse selection both work to raise the cost of the insurance for everyone. Because insurance payouts are higher or more frequent, everyone's premiums rise, to cover the cost to the insurer. If the premiums get too high, then some people in low-risk groups may simply leave the scheme. This will mean further increases in premiums for those that remain, which may prompt more flight from the scheme, and potentially the breakdown of the insurance scheme.

Adverse selection causes efficiency concerns for insurance generally, as does moral hazard. (Moral hazard is where the fact of being insured prompts people to take risks that they would otherwise strive to minimise. People tend to consume more insurance when they bear a smaller share of the costs that insurance covers. For example, suppose that once Careless Cara is fully insured against theft she takes less care with securing her possessions. Suppose that the overall risk of theft is 5%, but that calculation is based on people who take normal security precautions. Careless Cara gets a premium rate that reflects that 5% but her subsequent

Heath notes three main concerns raised against statistical discrimination in actuarial insurance. One objection is that "it is unfair to penalize individuals for circumstances that are outside of their control, or for things that are not their fault" (Heath, 2014, p356-7). The basic idea is that people should be treated in accordance with what they can be held responsible for, and not be treated differently over things they cannot control. From this perspective, it would be ok to charge more for car insurance to a person with a history of poor driving. However, it would not be ok to charge a young man with no history of poor driving more for his car insurance simply because he is a young man, and young men as a group are involved in more car crashes.

Heath rejects the 'outside your control' complaint on the grounds of inconsistency. He notes that it seems to subject private insurance organisations to a more demanding standard of ethical conduct than that of any other industry. "Individuals routinely benefit from their natural endowments (intelligence, beauty, creativity, etc.) or from brute luck (plentiful rain, an early frost, a change in interest rates, etc.) in market transactions, and we think nothing of it." (Heath, 2014, p358). So Heath is not disputing that we generally think it unfair to treat people on the basis of things they are not individually responsible for. Instead, he is saying that this unfairness is not uncommon in life generally, and if we don't condemn it when it involves other business agents, then why do so for insurers?

Another objection Heath canvasses is the concern that high-risk individuals are offered policies at such a high cost (Heath, 2014, p362). Heath suggests that insurers are not acting unfairly to ask high-risk individuals to pay a high price for their insurance. One reason he gives focusses on the purpose of private insurance, which is primarily risk management. Each member joining an insurance scheme is aiming to manage their own risk - the aim of taking on someone else's loss is not a reason anyone would buy into an insurance scheme

---

carelessness means that her risk is double that). Powell and Goldman (2015) studied the difference in health spending of employees who selected the most or least generous insurance plans at a large firm, and found that moral hazard explained 53% of the difference, with adverse selection explaining the remaining 47%.

(Heath, 2015, p360). Each member is asked to pay an amount that reflects what the insurer estimates the member is likely to claim back over the life of the policy, rather than an amount someone else is likely to claim. I suggest we think of this idea as 'compensatory contribution'. You are asked to contribute in proportion to the insurance company's expectation for compensating you in claims. Suppose that we remove a profit-focus from the equation; after all, insurers can be not-for-profit. In this case, the high-risk applicant is only being asked to pay in proportion to what the insurer expects the applicant will take out over the life of the policy. It does not seem obviously unfair that if the insurer has reasonable grounds to expect that a person will make a larger claim on the insurance, they should charge that person a larger amount to insure them.

The problem with high-risk individuals not being able to afford insurance is not simply about the cost of insurance but rather more fundamentally about the cost of their risk. It is not the insurer's fault if a person's risk is very costly, and it is not something the insurer has control over. Daniel Palmer (2007, p117) notes that the concern at issue here is that some people are vulnerable in terms of accessing certain goods or services they need, such as expensive health care – and that this is a social issue that goes well beyond insurance practices. We could complain about Heath's and Palmer's position on this, if we think that high-risk applicants may disproportionately have qualities that make them more vulnerable and needy. In the retail sector, sellers are supposed to have a special duty to their vulnerable customers (such as children or the confused elderly), which might mean foregoing a sale in order to protect the interests of their vulnerable customers. Suppose that for insurance, elderly people and people suffering from mental illness are likely to be high risk in terms of healthcare. If a significant proportion of high-risk applicants are likely to have such qualities, oughtn't insurers to have a special duty of care to their welfare, similar to the retail sector?

Recall (from my methodology in chapter one) that I am assuming the real-life context of individual decision-makers operating in a Western, developed country, under existing laws and competitive pressures. In this context, for parts of the population that lack the wherewithal to insure their risk, the duty of care ought not to fall on their insurer. The duty

instead falls on sections of the public sector that deal with members of society who have urgent needs for society's help. This is because the socialisation of the losses and life costs of those with urgent needs is what social welfare is for, and is the role of governments not insurers. As Deborah Stone notes, 'Insurance has a different purpose and targets a different population than welfare. The goal of insurance is to "offer security to individuals who have something to lose [e.g., a job, savings, earnings potential] rather than assistance to the needy, who have little or nothing to lose" ' (David Moss, cited in Stone, 2002, p59). The insurer simply offers the high-risk applicant a more-or-less beneficial insurance relationship (potentially with room for the terms of the relationship to improve based in part on the future behaviour of the applicant).

The other primary ethical concern with insurance addressed by Heath is that insurance is basically the same as "plain old-fashioned discrimination" (Heath, 2014, p351), because people end up being treated on the basis of group characteristics rather than as individuals, and this offends against our idea of equal treatment of individuals. Heath argues this is wrong, drawing on the point that I earlier labelled as 'compensatory contribution'. Heath notes that "there is no way to guarantee that the *ultimate benefit* of entering into an insurance scheme will be the same for all individuals (if there were, people could just save), [so] we must ensure that the *expected benefit* be the same for all." (my italics, Heath, 2014 p356). The expected benefit is exactly reflected in the calculation of the 'actuarially fair' premium for each individual, being a premium that will cover the expected loss from the individual's membership in the insurance scheme. So Heath argues that applicants are treated equally as individuals coming in to the insurance arrangement – they are treated equally on the basis of the expected benefit to them from the insurance arrangement.

I think Heath makes a significant point. However, his argument seems on shakier ground if we consider an admission from him earlier in his paper. Heath (2014, p345) writes that "In reality, insurers are often unable to determine the risks that each individual faces. Thus they use a system of more-or-less broad classification, in order to determine which "risk pool" or class an individual falls into." Daniel Palmer (2007, p114) too agrees that insurers must discriminate between individuals not really as individuals, but in terms of their falling into

108

certain categories of persons. In light of this, it seems harder to hold that the equality at the start of the insurance arrangement is for individuals rather than classes or groups. Alternatively, perhaps we need to consider further what it is for an organisation to treat someone as an individual.

Recall Kasper Lippert-Rasmussen's (2011, p54) definition for treating someone as an individual: "X treats Y as an individual if, and only if, X's treatment of Y is informed by all relevant information, statistical or non-statistical, reasonably available to X." In chapter three, I raised the concern that if there was solid non-statistical information available about a feature of an individual, then adding statistical information about that feature seemed only to detract from treating the person as an individual.[42] However, Lippert-Rasmussen's definition works better in the situation where required information is missing or insufficient. Lippert-Rasmussen (2011) notes that we never have all the possible information about an individual. Yet we need to make assessments of other people, so we must generalise from our incomplete information. In so doing, we may need to draw on knowledge and experience of other people. On this basis, Lippert-Rasmussen[43] (2011) argues that people cannot have a fundamental right to be treated *fully* as individuals, if this means judging people without reference to *any* group generalisations. I agree with Lippert-Rasmussen insofar as it concerns insurance organisations dealing with applicants. Insurers are looking years into the future; they may need to draw on knowledge of other people over time, to form a picture of the applicant in the future.[44]

Lippert-Rasmussen suggests that part of treating someone as an individual could involve using information that categorises them as part of a group, if we are also careful not to ignore other information that is particular to that individual. In the insurance context, for example, this could mean calculating a 'susceptibility to skin cancer' score for an individual

---

[42] For example, if X has a cholesterol test and scores very well on this, then drawing on information that 85% of people of X's age and sex score poorly with regard to cholesterol seems only to detract from treating X as an individual.

[43] Part of this paragraph is taken from Scholes (2016).

[44] Because insurers have access to information provided by actuarial modelling and assessment (American Academy of Actuaries, 2006), insurers can do this more objectively and accurately than the average person.

based on generalisation from groups, but also taking account that this individual is having yearly mole checks to facilitate early detection of any problem moles. This also fits Eidelson's 'character condition' (2013, p375), discussed in chapter three, that requires the decision-maker to give reasonable weight to evidence of the ways the subject has exercised her autonomy in giving shape to her life, where this evidence is reasonably available and relevant to the determination at hand. Overall, this means that individualised treatment can come in degrees, depending on how much individualised information the decision-maker takes into account. A decision can be less discriminatory the more it incorporates individualising information about an applicant, and the less it relies on statistical stereotype.

To get a sense of the 'discriminatoriness' of insurance, we can compare it with statistical discrimination in another applicant setting. Suppose that there is a housing shortage and a landlord is vetting applications from potential tenants. He has heard that people under 25 make worse tenants, and so discards applications from applicants that list their age as under 25. I suggest there are some morally relevant differences between the differential premium decisions made by insurers on the sorts of personal information that insurers gather, and the decisions of landlords discarding the tenancy applications of undergraduate students. The landlord makes salient just one or two pieces of information to judge an applicant; in doing so, the individuality of the applicant is very much ignored. The insurer, by contrast, may take a great many pieces of information about an applicant, and makes salient numerous pieces of information, to sort the applicant into a risk-class. In the case of insurance applicants, it could even be that the particular combination of detailed information an insurer considers for some applicants *is* individuating – that it does single out the applicant as an individual. Even if it does not though, my claim is that the insurer's treatment of the applicant is far more individualised than that of the landlord. The difference is one of degree. Applicants for insurance are treated *more* as individuals, despite being put into a 'risk pool' and having their premium determined according to the risk pool.

Another difference is the consequences of applying for the applicant. For tenancy seekers screened out due to statistical discrimination, the consequences are basically all negative.

110

They spend time preparing their applications and receive no useful feedback from the landlord. They do not even know, from their rejection, that they were not the best applicant (it is possible that they were a very prudent and responsible 24 year old, and they were better for the position than the others applying). In the insurance situation, there are likely to be some positive consequences from applying, even if the premium offered is too high for the applicant to afford. Insurance companies operating in a competitive environment are incentivised to assess risk as accurately as possible, in order to offer premiums that will be both attractive to customers and sustainable for their business. Insurance companies will have a much better overall picture of risks than applicants. Insurers will be asking questions relevant to risk that applicants may not even know to consider. The actuarial knowledge of insurers can benefit applicants through signaling information to them on their risk. Suppose that a couple of insurance companies quote you a high price for your health insurance, or an average price but with some significant exclusions. It would probably be to your benefit to get your health checked, to get the excluded factors examined, or to look at your overall health-maintaining efforts, and see what you might do to improve.

Finally, we can consider the knowledge and expectations of applicants engaging with the insurance industry and seeking insurance, in comparison with tenancy seekers. The latter probably expect that if they complete an application form, the information they supply will be considered in good faith by the landlord. Applicants may expect that if their tenancy history, credit history or referees do not recommend them to the landlord, their applications will be dismissed. But I submit that applicants will not expect their applications to be dismissed on the basis of statistics relating to being 24 rather than 26 years old. Insurance is different. Having an insurance industry allows applicants the opportunity to accept to be treated on the basis of some collective statistics and pay premiums in order to mitigate the financial risk posed by potential future adversity. Applicants for insurance are aware that the personal information they provide will form the basis for statistical discrimination. Applicants are, in a sense, choosing the discriminatory practices in order to gain the insurance benefits. Unlike the practice of the landlord, the knowledge and expectations of insurance applicants mean they can consent to the insurance application process.

In conclusion, I suggest that the statistical discrimination in insurance may be less 'discriminatory' than similar discrimination in other areas such as employment or the rental market. It is less discriminatory because it can treat applicants, comparatively, more as individuals. The discrimination is also less morally problematic because it can produce beneficial information for all applicants as part of the process; and because applicants, through their knowledge and expectations, have the opportunity to consent to it as part of the application process. [45]

4.3 The broader decision-making context: Risk spreading and resources

In this final section, I discuss a point from Frederick Schauer concerning "all-or-nothing" decisions, and relate this to the selection of applicants by organisational decision-makers. An "all-or-nothing" decision is a general feature of a particular decision-making context that Frederick Schauer argues tends to support the use of statistical discrimination. I will argue that whether a context is or becomes an "all-or-nothing" affair can depend on the resources available to the decision-maker to gain further information or to spread risk. I suggest that applicant selection decisions may differ from Schauer's examples in this respect, offering more opportunity for the decision-maker to spread risk.

---

[45] A concern here is that insurance may not always be transparent and open to voluntary consent as I have suggested. Carol Heimer (2002, p128), for example, points out that:

> As ties between insurers and other organisations have become stronger, insurance has become a necessity for participation in the core economic and social activities of contemporary life. What should not be overlooked here is the role of insurers as extralegal regulators. In requiring insurance coverage as a condition for operating a business, owning a home, driving a car, holding office, or engaging in any number of other activities, governments, employers, banks, and other organisations are also requiring policyholders to follow insurers' rules. But unlike government regulations, insurers' rules are not established after public debate periods for comment and revision, or votes by an elected body.

Another concern is the amount of power insurers and their agents have to determine questions such as who is liable, questions of causation and responsibility. "In deciding when and whether to defend and pay claims, liability insurance claims personnel regularly decide who or what caused the event or situation out of which the claim has emerged." (Baker, 2002, p43) These decisions are affected by incentives for others involved in the judgments. For example, in the US context, "HMO doctors were more likely than the doctors in private practice to diagnose an injury or illness as work related" (Baker, 2002, p43) because workers compensation paid HMO doctors more than the HMO did, while private health insurance paid more than workers compensation for doctors in private practice.

Schauer (2003, pp87-88) notes that in some of our everyday decision-making under uncertainty, we have a method for making rational judgements: we can use the principle of expected value to decide what to do. The principle of expected value says that we should take the probability of an outcome for an action and multiply it by the value of that outcome to us in order to decide whether to take the action. For example, suppose that you are in a second-hand store deciding whether to buy a piece of clothing that you are unable to try on due to a problem with the fitting rooms. You are a pretty good judge of whether clothes fit you: you would estimate this correctly three-quarters of the time. You would be willing to spend $20 on this item if it fit. You judge how much to pay by multiplying the value of the outcome to you ($20) by the probability of the outcome occurring (0.75), thus: $20 x 0.75 = $15. In this case, you should spend up to $15 on the item.

The principle of expected value is not useful if a decision is an 'all-or-nothing' affair. As Schauer (2003, p91) says, a sports umpire who is 75 percent sure a player has committed an infraction cannot send off a player for three quarters of the usual penalty time; police officers cannot do partial stops; and you have to decide whether to hire this babysitter or not. Note how this is unlike the insurer, who can draw on detailed information at her leisure and use scaling to match a probability to a level of insurance premium; low risk to low premium, medium risk to medium premium, and high risk to high premium. In theory, the insurer could offer some kind of insurance opportunity to any applicant with a risk-assessment in between zero percent and one hundred percent (hence excluding only those applicants with no risk of the event occurring, and those applicants where there is certainty the event will occur). For decision-makers in 'all-or-nothing' situations:

1) there may not be detailed information of their choice to assist in their assessment

2) there is no scaling opportunity to match risk to expected outcomes in the decision that must be made. If a customs officer thinks there is an 80% chance that someone is carrying contraband, it makes no sense to search 80% of the person.

The difference is between situations where we can spread a good or an opportunity in degrees, over many people, over many times, and those where we cannot. For example,

take a good such as a cake. We are not required to give the cake just to one person; we can give several people a percentage of the cake. Or, take an opportunity such as investing money: we do not need to put our investable pot of money all in one investment; we can spread it over many investments in proportion to the estimated risk and return. By contrast, take the opportunity of the weather. Suppose that there is a 20% chance of rain today, and you are making a decision about taking your umbrella. It does not make sense to try to take 20% of your umbrella with you, nor to take your umbrella for 20% of the day. It's an 'all-or-nothing' affair: either you take your whole umbrella for the whole day or you leave it behind.

At first glance, selection decisions seem to be clear cases of "all-or-nothing" affairs: as with the babysitter case, either you offer to insure or hire or admit the applicant, or you do not. As Schauer notes, in such situations, we are faced with an uncertainty but the principle of expected value is not useful. Yet we must still make a decision. Schauer suggests that the probabilities that statistics provide can inform this situation we face in 'all-or-nothing' decisions. While probabilities do not remove the uncertainty, neither does the principle of expected value; at least the statistical generalisations help inform the 'all-or-nothing' situation that cannot in practice be resolved with the principle of expected value. If many everyday decisions are such 'all-or-nothing' affairs, suggests Schauer, this is a reason to explain our use of generalisations to exclude 'riskier' individuals, or have them bear the full brunt of a disadvantage. The reason is that we are unable to tailor our decisions to the levels of risk.

I want to make two points in response to Schauer's insights. The first point is that some situations that are "all-or-nothing" affairs do allow decision-makers to gain more detailed information. This is a difference distinguishing the applicant selection decisions I am interested in. Some of the decisions Schauer identifies need to be made on the spot, such as the sports umpire sending off a player or the police officer stopping a car. In these cases, the decision maker has very little control over the timing or informational resources for making a decision. This is not the case with applicant selection decisions. The organisational decision-maker has plenty of time to deliberate over their choice. This is important because more time means a decision-maker can draw on more resources when deciding whether

and how to offer an opportunity to an applicant. For example, if insurance application assessors are deciding whether to offer insurance, or if an employer is deciding whether to keep an applicant in the applications process, they can choose to request extra information.

Furthermore, some decisions that initially seem to be "all-or-nothing" affairs can allow for the gathering of more detailed information, if the decisions are repeated rather than one-off. Insurance policies can be updated on a yearly basis, for example, taking into account new information from the previous year (Abbring, Chiappori & Zavadil, 2008, p.1). In this way, the organisation's assessment of the individual is broadened to include their experience of the individual over time. In insurance, this is referred to as 'experience rating'. This can include experience of the individual's behavior over this period; for example, insurance policies can offer a 'no-claims' bonus (Pinquet, 2012). An added benefit of experience rating in insurance (as opposed to class rating) is transparency – the rating basis is visible to the individual insured (Meyers, 1985). Alternatively, consider the selection decision of whether to hire the babysitter. This decision could easily be repeatable; for example, if we need a babysitter once a month, the decision "whether to hire babysitter X" can be revisited each time babysitting is required and revised in light of experience.

Nonetheless, it must be admitted that some decisions about selection of applicants will be one-off rather than repeated. In this case, I bring in my second point: that in some "all-or-nothing" affairs the target of the decision can be widened in a way that spreads risk. Whether this is so in a particular case will depend on what risk we want to spread and what resources we have available to do this. For example, if our concern is that the babysitter is not trustworthy, we can ask neighbours or family to call in while we are out, to check how things are going. If we are concerned that the babysitter lacks important skills, we could request the babysitter complete a first aid course. In these ways, we widen the target of the decision to include resources that can mitigate the risk. From an organisational perspective, this facilitates organisations to be able to treat applicants more as individuals. However, organisational decision-makers may need a certain level of resources to be able to engage in this.

As noted in chapter three, an organisation's responsibilities in regard to spreading risk will depend on the resources it has at its disposal. In the babysitter example in this paragraph, the required resources are social: having trusted neighbours and family who are willing to help out. In other cases, the necessary resources may be financial or cultural. For example, a large organisation may have the financial resources to deploy other staff temporarily to cover a vacancy while the decision-maker seeks more information on applicant risks. The organisation may also be able to support the decision-maker with Human Resources expertise while she does this. A small business, by contrast, may not have specialist HR expertise to hand, and may not have enough financial resources to cope with leaving a vacancy open for longer than necessary.

Conclusion

In this chapter I have argued that the purpose of the organisation in society is ethically significant to evaluating its use of statistical discrimination to assess individuals. Regarding courts, statistical discrimination as a method seems contrary to the ends of serving justice in particular cases, whereas in actuarial insurance it fits the purpose of the organisation. I have also argued that the 'discriminatoriness' of the use of risk-based statistical discrimination can vary depending on how detailed the examination of the applicant is: more detailed is less 'discriminatory'. Moral problems are further lessened by the provision of benefits, and the transparency of the organisation's use of the method. This may be the case with actuarial insurance; however, whether this is likely to occur in other applicant selection situations is an open question. Finally, I argued that a decision-making context that may prompt organisational decision-makers to use statistical discrimination to deal with uncertainty ("all-or-nothing" decisions) can change depending on the information the decision-maker can access and the resources they have to spread risk. This can affect the opportunity for an organisation to treat applicants as individuals.

This chapter presented two examples that I think fall at the extremes in terms of our expectations for individualised judgement of individuals (court cases) or statistical judgement of individuals (insurance). It would be fair enough to wonder about the relevance of the analysis to assessing statistical discrimination regarding applicants in other selection

situations. This is particularly the case if aspects of the three dimensions noted above differ significantly in different situations; for example, if the purpose of the organisation is contested, or if the organisation aims to benefit 'risky' applicants (rather than exclude them) through its use of statistical discrimination. I will address this concern in my next two chapters. Chapter 5 offers an in-depth discussion of statistical discrimination in the treatment of 'risky' students in a higher education context, focusing on the use of analytics to identify and treat differently those students at higher risk of failing or dropping out. Chapter 6 presents an in-depth discussion of statistical discrimination in hiring. In these chapters I hope to show the ethical significance of the dimensions identified in this chapter to evaluating statistical discrimination, and include some broader considerations on risk statistics and treating applicants as individuals.

Chapter Four References

Abbring, Jaap, Chiappori, Pierre-Andre & Zavadil, Tibor (2008) Better safe than sorry? Ex ante and ex post moral hazard in dynamic insurance data. Tinbergen Institute Discussion Paper, TI 2008-075/3. Retrieved 28 October 2016 from http://dare.ubvu.vu.nl/bitstream/handle/1871/12965/08075.pdf?sequence=1

Allen, Ronald (2015) A note to my philosophical friends about expertise and legal systems. *Humana.Mente Journal of Philosophical Studies*, 28: 71-86.

American Academy of Actuaries (2006) The roles of the actuary in the selection and application of actuarial models: concepts on professionalism. Discussion paper no 7. Retrieved September 2106 from http://www.actuary.org/files/publications/models_june06.pdf

Arthur, Jeffrey B (1994) Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal, 37* (2): 670-687.

Avraham, Ronen, Logue, Kyle D and Schwarcz, Daniel (2013) Understanding insurance anti-discrimination laws. *Public law and Legal theory Research Paper series, paper no. 289, Law & Economics Research Paper Series, paper no. 12-017*: 1-52.

Baker, Tom (2002) Risk, insurance, and the social construction of responsibility. In Tom Baker and Jonathan Simon (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (Pp33-51). Chicago and London: University of Chicago Press.

Baumle, Amanda K. and Fossett, Mark (2005) Statistical Discrimination in Employment: Its Practice, Conceptualization, and Implications for Public Policy. *The American Behavioral Scientist,* 49: 1250–1274.

Bunn William B. III; Stave, Gregg M.; Downs, Kristen E.; Alvir, Jose; Dirani, Riad (2006) Effects of smoking status on productivity loss. *Journal of Occupational and Environmental Medicine,* 48 (10): 1099-1108. DOI: 10.1097/01.jom.0000243406.08419.74 Retrieved from http://journals.lww.com/joem/Abstract/2006/10000/Effect_of_Smoking_Status_on_Productivity_Loss.18.aspx

Canadian Superior Courts Judges Association (n.d.) The role of the courts. Retrieved from http://www.cscja-acjcs.ca/role_of_courts-en.asp?l=4

Courts of New Zealand (n.d.) The role of the courts. Retrieved from https://www.courtsofnz.govt.nz/about/system/role/overview

Cragg, Wesley (1992) *The Practice of Punishment: Toward a Theory of Restorative Justice*. Routledge, London and New York.

Enoch, David & Fisher, Talia (2015) Sense and "Sensitivity": Epistemic and instrumental approaches to statistical evidence. *Stanford Law Review*, 67 (3): 557-611.

Federal Trade Commission (July 24, 2007). FTC releases report on effects of credit-based insurance scores. Retrieved 26 July 2016 from https://www.ftc.gov/news-events/press-releases/2007/07/ftc-releases-report-effects-credit-based-insurance-scores

Guerra-Pujol, Enrique (2014) Visualizing probabilistic proof, 7 *Washington University Jurisprudence Review* 39. Available at http://openscholarship.wustl.edu/law_jurisprudence/vol7/iss1/6

Hallevy, Gabriel (2013) *The right to be punished: Modern doctrinal sentencing*, DOI 10.1007/978-3-642-32388-1_2, Berlin Heidelberg: SpringerVerlag

Heath, Joseph (2014) Reasonable restrictions on underwriting. In *Morality, competition and the firm: The market failures approach to business ethics* (Pp.345-372) Oxford: Oxford University Press.

Heimer, Carol A (2002) Insuring more, ensuring less: The costs and benefits of private regulation through insurance. In Tom Baker and Jonathan Simon (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (Pp116-145) Chicago and London: University of Chicago Press.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *Journal of Political Philosophy,* 15 (4): 385–403. DOI: 10.1111/j.1467-9760.2007.00285.x

Lord Low of Dalston (16 May 2007, 4.45pm) Offender Management Bill. Lords Hansard Home Page. Parliamentary business, Publications and Records. Retrieved from http://www.publications.parliament.uk/pa/ld200607/ldhansrd/text/70516-0004.htm

Meyers, Glenn (1985) An analysis of experience rating. PCAS. Retrieved 28 October 2016 from http://www.casact.org/pubs/proceed/proceed85/85278.pdf

Palmer, Daniel E (2007) Insurance, risk assessment and fairness: An ethical analysis. Insurance ethics for a more ethical world, *Research in Ethical Issues in Organisations*, 7: 113-126.

Pinquet, Jean (2012) Experience rating in non-life insurance. Ecole Polytechnique, Center National de la Recharche Scientifique, Cahier no 2012-10. Retrieved 28 Ocrtober 2016 from https://hal.archives-ouvertes.fr/file/index/docid/677100/filename/cahier_de_recherche_2010-12.pdf

Powell, David and Goldman, Dana (2016) Disentangling Moral Hazard and Adverse Selection in Private Health Insurance. NBER Working Paper N0. 21858. Accessed 29 June 2016: http://www.nber.org/papers/w21858

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Scholes, Vanessa (2014) You are not worth the risk: Lawful discrimination in hiring. *Rationality, Markets and Morals*, 5: 13-29.

Scholes, Vanessa (2016) The ethics of using learning analytics to categorize students on risk. *Educational Technology Research & Development.* DOI: 10.1007/s11423-016-9458-1

Stone, Deborah (2002) Beyond moral hazard: Insurance as moral opportunity. In Tom Baker and Jonathan Simons (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (Pp52-79). Chicago and London: University of Chicago Press.

Tyler, Tom R (2014) How do the courts create popular legitimacy?: The role of establishing the truth, punishing justly, and/or acting through just procedures. *Faculty Scholarship Series.* Paper 4991. Retrieved from http://digitalcommons.law.yale.edu/fss_papers/4991

Chapter 5   Using learning analytics to risk-screen applicants to tertiary education[46]

Introduction

My aim in this chapter is to show the relevance of the concepts and conclusions from the earlier chapters to analysing statistical discrimination in a real life context. I will examine statistical discrimination in organisational selection of applicants for admission to tertiary education. Statistical discrimination in this context will occur through using analytics to risk screen students. In my analysis I will bring in considerations of the purpose of higher education institutions, the benefits to students of risk screening, the lack of respect for students as individual agents, and the importance of transparency. There are good reasons for higher education institutions to use learning analytics to risk screen students. Analytics can better predict which students are at greater risk of dropping out or failing; institutions can then use the statistics to treat 'risky' students differently. In the first part of this chapter, I explain learning analytics and the impetus for higher education institutions to use the analytics information.

The second part of this chapter explores the literature on the ethics of using learning analytics. Ethical concerns raised by the deployment and intended deployment of learning analytics have recently been the subject of increasing discussion (see, e.g., Campbell, deBlois & Oblinger, 2007; Simpson, 2009; Kay, Korn & Oppenheim, 2012; Slade & Prinsloo, 2013; Johnson, 2014; Pardo & Siemens, 2014; Polonetsky & Tene, 2014; Prinsloo & Slade, 2014; Sclater, 2014a; Willis, 2014; Willis & Pistilli, 2014; Sclater & Bailey, 2015). These discussions tend to take existing ethical and legal principles on the use of data, such as transparency, consent, choice, accountability, privacy and security, and translate them to the education analytics context. There is less discussion of the ethics of subjecting an individual to intervention on the basis of information about group risk. This chapter contributes to this topic, drawing on philosophical accounts of wrongful discrimination to

---

[46] This chapter is a very slightly modified version of a paper accepted for publication in *Educational Technology Research and Development* in June 2016 and published online September 2016. See Scholes (2016). The published version is available at http://link.springer.com/article/10.1007/s11423-016-9458-1

examine in what sense intervention on the basis of learning analytics might be ethically concerning.

Use of analytics to risk screen students is not commonly viewed as discrimination. I address this in the third part of this chapter. I apply some of the philosophical accounts of discrimination from chapters one and two to analyse this use of learning analytics as a form of discrimination. As noted in chapter four, an organisation's purpose can influence the morality of its use of statistical discrimination. Education institutions have a different purpose from other organisations discussed in earlier chapters, such as insurance organisations, rental agencies or employers. Notably, institutions of higher education typically aim to identify 'risky' students in order to direct more resources to them, rather than to disadvantage or exclude them. Nonetheless, some students identified by education institutions as 'risky' are likely to be disadvantaged by this, and some may be excluded. My analysis suggests the principal ethical concern with the differing treatment of students is a comparative failure to recognise students as individuals, which may affect students as agents.

In chapter three, I analysed an argument from Frederick Schauer (2003) on individual assessment. In the fourth part of this chapter I draw on that analysis, translating it into a learning analytics context. Schauer takes an instance of what we think of as 'individualised' evidence, namely, testing of individuals, and argues that this is based on group generalisations. This is because the tests themselves are normed using statistics from the groups that sat the test. If 'individualised' evidence is grounded in statistics about groups, then in this respect, says Schauer, such judgements are not different in kind from judgements based directly on group risk statistics. I compare the treatment of a student under learning-analytical risk screening with a more traditional individualised assessment, namely an individual examination. I find the former ethically wanting in terms of respect for students as individual agents. However, as noted in chapter four, other dimensions of the decision context such as benefits to the individuals affected and transparency can affect the morality of risk-based statistical discrimination. It seems learning analytics can bring

important benefits, and that education institutions will use them, hence my question: are there ways of using learning analytics that better respect students as individual agents?

In the fifth, final part of this chapter, I suggest some educational design focusses for learning analytics that offer ways to mitigate the concern about treating students as individual agents. I make three main recommendations to support recognition of and assessment of students as individual agents. The first two recommendations are designing features into courses that promote greater use of effort based factors, and greater use of dynamic rather than static risk factors. These recommendations are drawn from my analyses in chapter three. The third recommendation is greater use of sets of statistics specific to individuals. This recommendation draws from a point raised in chapter four which I develop further in this chapter. I conclude with a brief note on how the power asymmetry between higher education institutions and students may influence how the risk of student dropout or failure is dealt with, based on a point raised in chapter three.

5.1 Risk-focussed learning analytics

Institutions offering higher education are interested in identifying students at greater risk of failing or not completing their course of study, and intervening before this happens. To facilitate this goal, institutions can use learning analytics: statistical analyses of data gathered on students to better support desired education outcomes for students, as individuals and as groups. The Open University, for example, describes its learning analytics practice as using 'raw and analysed student data to proactively identify interventions which aim to support students in achieving their study goals. Such interventions may be designed to support individual students and/or the entire cohort' (Open University, Sep 2014, p.1). Computer tracking and analysis of large amounts of data from all students can identify factors statistically correlated with worse performance, especially failing or dropping out. Predictive risk modelling can be used to assign a risk categorisation to individual students (see e.g. OAAI, 2012; Jia, 2014; Open University, Oct 2014; Wagner & Hartman, 2013). Institutions can then treat 'high risk' individuals differently, with the aim of ensuring they do not end up counting as negative statistics for completion.

There is no single agreed definition of 'learning analytics' used by education researchers and institutions. Long and Siemens (2011) proposed restricting the term 'learning analytics' to use of data focussed on the learning process, as opposed to data analysis for business intelligence at the institutional level. It is worth noting Sclater's 2014 report on learning analytics in the UK, which suggests institutions see learning analytics data more broadly as a resource for various stakeholders 'from individual students and their tutors to educational researchers, to unit heads and to senior management' (Sclater 2014b, p.4). However, I am more interested in the learning process and learning environment than in, for example, use of analytics to target fundraising efforts (see Campbell, deBlois & Oblinger, 2007). Accordingly, this chapter uses a learner-focussed definition of learning analytics from Slade and Prinsloo (2013). Learning analytics is '…the collection, analysis, use and appropriate dissemination of student-generated, actionable data with the purpose of creating appropriate cognitive, administrative and effective support for learners' (Slade & Prinsloo, 2013, p.4). This will include both demographic data and course engagement data from students.

Implementing learning analytics requires systematic collection and analysis of large amounts of data to identify risk factors across students. Students are then classified as a member of risk-bearing groups or categories (see Jia, 2014), and 'risky' students are subjected to different treatment (interventions). Outside the education context, this process is not uncommon. Applicants for a bank loan, or for insurance, are screened, and the organisation forms some kind of risk profile for the applicant. In the higher education context, the practice of learning analytics is typically expected to involve screening students on group risk factors for the purposes of designing and targeting interventions. The risk factors might include, for example, part-time status, gender, ethnicity, nationality, number of years of prior education, highest level of educational qualification, student engagement with courses, and accessing library resources (see, e.g. Jia, 2014; Open University, Sep 2014). Bichsel (2012) reports that Purdue University's learning analytics system uses grade information, demographics, and existing data on student engagement. Sclater (2014b)

reports both Nottingham Trent University and the Open University have identified accessing library resources as a predictor of student success.

Interventions for 'risky' students could include restrictions on the level of course studied, restriction on the number of courses studied, or imposition of prerequisites, which might mean a requirement to take extra courses (such as a bridging course). Interventions could also simply involve offers of additional support, such as extra phone calls to encourage engagement, or referrals to academic (or other) support services. However, even these less restrictive interventions may direct 'risky' students to do more work, or impress on them an expectation that they do more work. Interventions may be appropriate for some of the group of 'risky' students, but there may be individuals in the group who would be better off without the intervention (Sclater, 2014b), in which case it is a burden on them. An institution could also try to discourage a student from continuing with study (see Open University Sep 2014, p9), or avoid putting resources into a student the institution thinks is not going to succeed (Bichsel, 2012).

Using analytics to identify, and intervene with, students at greater risk of not completing their course of study is increasingly seen as an economic and pedagogical imperative for education institutions (Kay, Korn & Oppenheimer, 2012; Slade & Prinsloo, 2013). This will be particularly acute for distance education institutions offering courses online (Prinsloo & Slade 2014). Online distance education learners continue to have poorer retention rates than class-based learners (Lokken & Mullins, 2015), and online education provides huge opportunities for computer tracking and analysis of data on students' online behaviour. Together, these factors suggest institutions that provide online education have strong motivations to use group risk analytics to screen and categorise students. Proposed benefits of learning analytics include increases in student performance and student retention (Bichsel, 2012; Sclater, 2014b), through highlighting patterns of success or disengagement in students (Oblinger, 2012) that can '[inform] the design of more effective, appropriate and, importantly, more cost-effective student support' (Prinsloo & Slade 2014, n.p.). Importantly, the risk statistics and predictions for students generated through the use of learning

analytics can provide more information resulting in more systematic and effective interventions (see Sclater, 2014b).

5.2 Conceptions of the ethics of learning analytics

Theorists discussing the newly emerging topic of the ethics of learning analytics have looked to the ethical principles established for existing practices that may be related to learning analytics practices. For instance, as learning analytics involves gathering and analysing information from students, it could be seen as a type of research on students. Hence, some researchers have considered whether analytics has a normative basis in the principles of research ethics (see e.g. Pardo & Siemens, 2014; Kay, Korn & Oppenheim, 2012). However, Kay, Korn and Oppenheim (2012) note several differences between the contexts for research and for analytics that render the 'research ethics' conception of the principles less relevant. For example, 'consent' in a research context is specific to a particular research project being undertaken, and the intent of the use of the data in research is 'directed toward an agreed outcome' (Kay et al., 2012, p.22). Use of learning analytics does not typically involve such discrete, well-defined research projects. This does not invalidate consent as an ethical principle for learning analytics; instead, it suggests the practices constituting acceptable consent in research ethics do not easily transfer to a learning analytics context.

The desired context for learning analytics is continual data collection and processing for ongoing educational intervention and innovation. Unsurprisingly then, education theorists have conceived of the ethics of learning analytics in terms of the *institutional collection of data* from individuals. Slade and Prinsloo (2013), for example, identify three broad areas of ethical issues for learning analytics: the location and interpretation of data; informed consent, privacy and the de-identification of data; and the management, classification and storage of data. Sclater's comprehensive review of the literature on legal and ethical issues for learning analytics identifies the top four ethical foci as: "transparency, clarity, respect for users and user control" and notes that "[c]onsent, accountability and access also feature prominently" (Sclater, 2014a, p.59). Pardo and Siemens (2014) start with a research ethics base and focus on privacy and personalisation in learning analytics, yet even their four

principles of transparency, student control over the data, right of access and security, and accountability and assessment largely fit under the data collection concerns. Most of these concerns echo recommendations from legal guidelines for data protection and privacy (see, e.g. the Information Commissioner's Office's Data protection principles, n.d.). As these guidelines embed ethical principles for practitioners, so the focus in the current literature on transparency, consent, choice, accountability, privacy and security of data is well placed for guiding institutions' considerations of the ethics of learning analytics.

Implementing an extensive learning analytics system is not just another use of data with privacy implications, however. At the heart of the practice is the concept of categorising students according to the statistical risk that can be attached to them, and subjecting students to different treatment on this basis. This has given rise to some additional ethical concerns for the practice of learning analytics: discrimination, identity and the agency of students. The spectre of discrimination is raised by Polonetsky and Tene (2014); Sclater and Bailey (2015) and Slade and Prinsloo (2013). Identity and agency concerns are raised by Johnson (2014), Prinsloo and Slade (2014), Sclater (2014a) and Slade and Prinsloo (2013); and are addressed by the Open University (OU, Sep 2014). However, in comparison with, for example, the ethics of big data (see e.g. Crawford & Schultz, 2013; IWGDPT, 2014; Richards & King, 2014) there is less focus on these concerns and how they interact. Moreover, the recommendations regarding agency tend to be simply to inform students that learning analytics are used and get their consent. This chapter contributes to the literature by employing normative theories of discrimination to analyse the use of learning analytics. The chapter draws links to concerns for students as individual agents, and suggests how educational design can help support recognition of students as individual agents in learning-analytical assessment.

5.3 Discrimination and learning analytics

Discussions of discrimination in the literature on the ethics of learning analytics have focused on the concern that students may be stereotyped and mistreated in some way. This includes the concern that students may be at risk of being subject to prejudiced interactions

with the staff of the institution. The Open University's policy on ethical use of learning analytics warns that 'Analysis based on the characteristics of individual students at the start of their study must not be used to limit the University's or the students' expectations of what they can achieve' (Open University, Sep 2014, p.8). Differential treatment of individuals on the basis of groups they belong to may be considered discrimination when individuals are disadvantaged in some way. For interventions based on learning analytics, the disadvantage could include being required or encouraged to engage in unnecessary extra work that could also be an added expense (for example, being directed to take a bridging course).

As noted in chapter one, philosophers have not reached a consensus on how to specify the moral wrong(s) of discrimination. Philosophers who locate the moral wrong of discrimination in the false beliefs or hostile attitudes of the agents (Alexander, 1992; Arneson, 2006; Vallentyne, 2006) are unlikely to find anything of moral salience in the practice of learning analytics, as it does not feature attitudes or beliefs about the moral worth of the groups targeted. Suppose students are identified on the basis of a set of group characteristics correlated with completion risk such as gender, a certain level of prior education, and ethnicity, and are then treated differently. Neither the process of identification by the institution nor the differing treatment is likely to be based in beliefs or attitudes that members of these groups are morally inferior, nor motivated by hostility. So these agent based accounts of discrimination would see learning analytics as ethically unproblematic.

Deborah Hellman (2008) argues that the moral wrong in discrimination is not about the beliefs or attitudes of the agent, but instead whether the discriminatory act demeans someone, expressing a disregard for his or her moral equality. The social context gives the action its demeaning quality, along with the agent being in a position of power relative to the person subject to the discrimination. Tertiary education institutions are certainly in a position of power relative to students. But under Hellman's definition, use of learning analytics is only morally wrong if it brings into question the moral equality of those treated

differently, in a way that demeans them. This is clearly not the intent of the institutions using learning analytics, and it is at least unobvious that the social context renders this demeaning.

Kasper Lippert-Rasmussen's (2006) consequentialist position on discrimination suggests we should consider the consequences for the total amount of welfare, and distribution of welfare, in our assessment of differential treatment due to learning analytics. In fact, the distribution of the benefits of learning analytics is a key feature distinguishing it from risk screening in other areas such as lending and insurance. In these commercial settings, decision-makers want to identify applicants belonging to riskier groups to *exclude* them from a good, or give them worse conditions to access it. So an insurer might deny insurance or impose higher premiums, and a bank might deny loans or impose more restrictions on loans. In contrast, institutions using learning analytics aim to *benefit* the groups of students who are identified as bigger risks (for failing or dropping out), targeting them for interventions and extra resources. If so, learning analytics will produce good consequences. Moreover, it may produce a higher proportion of good consequences for worse off students, if students identified as bigger risks for failing or dropping out are unlikely to have been financially and socially well-supported for education, and more likely to have suffered from general welfare disadvantages. This may promote a better distribution of good consequences.

One scenario in which our expectations of good consequences from learning analytics fails is if institutions use learning analytics to identify high risk students to exclude them from a program of study, to exclude them from extra academic support, or even to exclude them from the institution. Bichsel (2012) reports the suggestion from an institutional research professional that

> […] it is very expensive to bring a student to campus who's not going to succeed, and so you have to be doing things all the way along the line to make sure

they're going to succeed *or recognise early enough before you've made a big investment that they're not going to succeed*. (my italics, Bichsel, 2012, p.15).

Note that the definition of learning analytics offered earlier is not compatible with continuing the enrolment of students while excluding them from academic support, given the purpose of learning analytics is to create 'appropriate cognitive, administrative and effective support for learners' (Slade & Prinsloo, 2013, p.4). However, it could be compatible with trying to exclude students from enrolling with the institution. Suppose learning analytics can deliver a very accurate prediction of a very high likelihood of failure for some students. Higher education is an expense for students, both in terms of the fees paid and the opportunities missed through the time spent studying. From a consequentialist perspective, we must ask whether it could be to the overall benefit of particularly high risk students, and to the institution, to exclude them from higher education study at a particular point in time. Certainly, not attending could be better for such high risk students, as a group, than to be admitted but offered no extra academic support, with many of the group then dropping out (Woodley & Simpson 2014, p474).

It is concerning to think of oneself being perceived and treated as a 'costly risk' by institutions due to one's group membership. This concern is brought out in Sophia Moreau's (2010, 2013) definition of wrongful discrimination as differential treatment of individuals that affects an individual's 'deliberative freedoms'. Recall that a deliberative freedom is a capacity to access core life opportunities without the concern that certain socially recognised traits will be counted as costs against oneself. Consider, for example, the non-White tenancy applicant who is more likely to be told "the place is already let" once the letting agent sees the colour of their skin. Relevant traits may also include gender or religious tradition, or any other traits that we think decision-makers ought not to mark negatively in distributing opportunities to access some goods and services. Moreau (2010) writes that '….[wrongful] discrimination involves a violation of someone's right to a certain deliberative freedom [….] denying them the chance to make a choice without having to think about this trait of theirs in the given context' (p.138) Discrimination thus affects liberty and autonomy.

In the higher education context, if students know they will be treated differently based on risks attached to certain traits they have, such as ethnicity or gender, this would potentially affect their deliberative freedoms, and thus their liberty and autonomy. In one sense, education institutions *are* marking students negatively, insofar as they are drawing on trait statistics to identify students as 'high risk'. However, unlike decision-makers assessing applications for tenancy or bank loans or insurance, decision-makers using learning analytics to mark students as 'high risk' are not typically trying to exclude or disadvantage identified students. A core part of the purpose of a higher education institution is to educate students, requiring students to succeed in passing courses and completing programmes. Institutions use analytics to intervene to try to foster student success. Some interventions may negatively affect the liberty and autonomy of some of the 'high risk' group; however, it is expected that interventions will benefit the majority of the group. Moreover, higher education intends to enhance a student's autonomy; if institutions implement learning analytics to do a better job of educating students, this aligns with their purpose and should further enhance students' autonomy. So it seems there is a potential trade-off between, on the one hand, recognition and respect for the autonomy of each student, and on the other hand, benevolent (if perhaps paternalistic) efforts to increase autonomy for groups of students.

Finally, Benjamin Eidelson (2013) suggests part of the badness of discrimination is simply the failure to treat a person properly as an individual, '... by failing to treat him as in part a product of his own past efforts at self-creation, and as an autonomous agent whose future choices are his own to make' (p.395). If a decision-maker treats a student differently just on the basis of their membership in a particular group, the decision-maker takes no account of how the student differs from other students in the group. This also entails that the decision-maker selects which feature of the student (namely, the feature of belonging to that particular group) will predominate in the decision-maker's assessment of the student. This is action on the basis of stereotyping, and stereotyping is a prime example of *not* treating a person as an individual. Eidelson's position does not sit easily with the use of learning

analytics based on the characteristics of groups, a use that takes little account of students as individuals. Assessing students based on group risk statistics, and treating 'risky' students differently, does appear to count as discrimination on this account. This is an ethical concern if we value treating people as individuals (at least to some extent), rather than stereotypes.

To conclude, using learning analytics to treat some groups of students differently is importantly different from other discrimination. It does not stem from hostile attitudes or false beliefs about the worth of the groups identified, and is not likely to demean the groups identified (not in the way that we think other discrimination demeans people). It is also not obvious that it is typically likely to negatively affect the freedom to deliberate about one's core life opportunities, or promote bad consequences for a socially salient group, in the way other discriminatory actions might. Indeed, learning analytics aims to produce benefits for the groups it identifies as at risk of poorer outcomes. Moreover, the interventions may enhance the autonomy of the majority of 'risky' students in the longer term, and using learning analytics to do this fits with the purpose of higher education. The one ethical concern from the accounts of wrongful discrimination that *does* characterise learning analytics in higher education is the failure to recognise a person as an individual, and any attendant effects on students as individual agents. This will be explored in the next section of the chapter.

5.4 Group risk statistics and individual students

The idea that people have an ethical claim to be recognised or treated as particular individuals, rather than merely on the basis of statistical generalisations about groups, is intuitively appealing. However, as discussed in chapter three, it is challenged by an argument put forward by Frederick Schauer (2003). As outlined in chapter three, Schauer (2003) uses an example of treating pit bull dogs differently due to breed statistics showing elevated risk of violence, to argue that what we think of as 'individualised' assessments really just reflect statistical generalisations from groups. Here I rework Schauer's example for the higher education context. The basic question is what a decision-maker in an education institution should make of any statistical evidence that some student groups

present significantly more 'risk' in terms of failure and dropout than others. Should a decision-maker arrange to have students examined individually and assessed against a certain threshold for likelihood of failure and dropout, and require interventions with extra work only for those who exceed the threshold? Or should a decision-maker simply sort students on the basis of group risk statistics and require intervention with extra work for all the 'statistically risky' students, without examination of them as individuals?

Suppose that institution A uses six characteristics to assess whether a student is a 'high risk' for failing or dropping out, namely gender, nationality, school socio-economic status, years of education, part-time status and highest prior qualification. Institution B instead gives each individual student some kind of pre-entry examination to assess whether they are likely to pass a program. Institution B's method is a more 'individualised' assessment – it focusses squarely on what the individual has done and is capable of. However, Schauer (2003) would argue that assessing individual students against an examination norm also involves assessing the individual students against a group generalisation; namely, a generalisation from the groups of students that, over time, were used as the basis for developing the norm. His argument would suggest that using individual examination to predict an individual's failure or dropout is not fundamentally different from using group risk statistics as a predictor of an individual's failure or dropout.

Schauer (2003) would also note that errors in prediction can still be expected with both institutions' methods. Institution A's six broad characteristics will be an inaccurate predictor. Some students with these characteristics, given the chance, would succeed as well as other 'less risky' students without interventions requiring them to do extra work. But examination procedures are imperfect too, and so will also inaccurately predict for some cases: some students who pass an examination procedure will not prove adequate to the program of study. Why should an institution use the results of an examination to intervene and impose extra work on a student, rather than using the fact that the student possesses a particular set of characteristics that means they fit a 'risky' profile?

Schauer's analysis thus suggests that in terms of generalising, an institutional decision-maker using membership in some 'risky' groups to make a judgment of a particular student does nothing fundamentally different from an institutional decision-maker who does an individual examination of a particular student. However, as I note in chapter three, this analysis misses something. Suppose that the following image is of a particular student, "Pat", who has the six characteristics that put him in a 'high risk' category.

[Image: Biagio D'Antonio. Samuel H Kress Collection, 1939.1.179 Retrieved from https://images.nga.gov]

Suppose too that the next image, reproduced multiple times, represents the category of high risk students.

[Image: Jan Steen, Widener Collection. 1942.9.81 Retrieved from https://images.nga.gov]

Finally, suppose that this image represents an examination norm:

Schauer's alternatives seem to be presented as: 'Pat put in the category of high risk students' versus 'Pat facing individual examination'. This makes it look like we are *considering Pat and the category of high risk students*, thus:
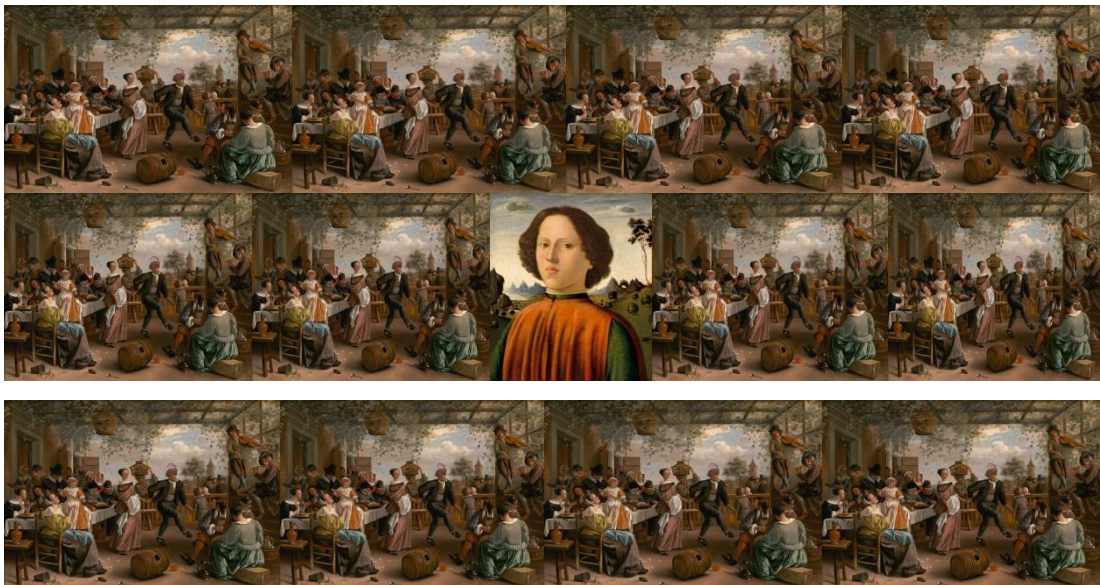

Fig.1 Considering Pat against the class of high risk students

versus *considering Pat and an examination norm, thus*:

Fig 2. Considering Pat and an examinations norm

On Schauer's view, if using a categorisation of high risk students and using an examination are both generalising at base, then we could say the institutional decision-maker's judgment of Pat is based on generalisations either way. However, the initial conceptualisation in Figure 1 is misleading, in the same way it was misleading when I discussed the example of 'Nobbo' the pit bull in chapter three. Figure 1 misrepresents the situation. If the decision-maker, when considering imposing an intervention requiring extra work, uses a classifier of the six characteristics as indicating a higher risk of failure or dropout, *Pat* the student is not judged. This is how Figure 1 should look:

Fig 3. Institutional decision-maker intervening based on group risk statistics [silhouette image credit shokunin, clipart. Retrieved from http://cliparts101.com/free_clipart/25163/student1]

Notice the placement of a silhouette of a student rather than an image of Pat. This is to illustrate that what the decision-maker is dealing with is a stereotype, not *Pat*, or *Pat's* academic ability. An examination, by contrast, includes participation by Pat and shows the individual abilities of Pat. As with the example of the pit bull 'Nobbo' in chapter three, the key point is that beyond being put in the 'high risk student' category, *Pat* the student is not considered.

As noted in chapter three, claims of group risk probability *extend over* an individual but there is a sense in which they are not *about* the individual qua individual. While institutional decision-makers do apply probabilistic claims to individual students, this does not mean they are picking out a feature of the individual rather than a feature of a group. An institutional decision using group risk statistics to impose interventions requiring extra work would be *applied* to individual students without really being *about the particular student*. An examination, in comparison, does consider the particular student that is Pat, and is about Pat. Judging individuals on the basis of group risk statistics does differ from the examination-based assessments, and the difference involves a comparative failure to deal with students as *individuals*.

When decision-makers in higher education institutions fail to consider the student they are dealing with as an individual, their decisions will feature only their own participation (or that of the council of the institution on whose behalf they act). There is little sense in which the individual student is recognised in the decision. The institutional decision-makers get to spread their 'selves' over actions that affect individual students, without the inclusion of students' 'selves'. This inhibits the extent to which the students being with dealt are included as agents in the process. This can also mean the decision unilaterally represents the institutional decision-maker's perspective and interests. Furthermore, in those cases where individual students suffer extra work unnecessarily due to institutional use of risk profiling, this may seem unfair.

Higher education institutions, however, have good reason to use learning analytics to make judgments about students. As noted initially, use of learning analytics has the potential to make decision-makers' judgments for targeting student support more effective and efficient, with benefits to students and the institution. Moreover Frederick Schauer (2003) argues that even if a practice of using group generalisations to judge individuals that is overall beneficial results in unfair decisions for some individuals, it can still be a good practice, and that it can be wrong to try to fix the unjust decisions. Schauer argues that whoever is deciding which unfair results to correct and how to correct them may be less reliable on average than letting a generalisation hold, with its inevitable instances of unfair judgments. Basically, even expert judges make mistakes when assessing individual information, and they might make things worse.

In the education context, this point receives support from a study by Vuong, Nixon and Towle (2011) on teachers' use of Cognitive Tutor, an electronic program of mathematics problems that provides feedback to individual students based on how they deal with the problems. The problems are organised in a sequence of modules, in accordance with research on appropriate mathematics pedagogy. Students must demonstrate mastery of the problems in one module before they can move to the next module; hence each module is a prerequisite for the next. However, teachers (our expert judges) are able to decide whether it would be better for their students if modules were skipped or taken in a different order, and can bypass the program to effect this. As it happens, 85% of teachers did this. As noted by Baker and Siemens (2014) the results of the research suggested that students were disadvantaged when teachers used their professional judgment to intervene and bypass the Cognitive Tutor sequencing. This supports Schauer's general point that allowing individual experts to intervene in the use of a system may not improve it, even when the experts have access to the relevant risk statistics.

In view of the foregoing, why should institutions do anything other than judge students on the basis of the probabilities attached to certain groups of which the student is a member? One reason lies in the value we place on assessing people as individual agents. As noted in chapter three, Edmonds (2006) argues that our judgments of the moral acceptability of the

discrimination from group generalisations hinge on an interplay of our assessments of causation and our assessments of responsibility or desert. We particularly value assessing people based on what they do – on their choices, efforts and achievements. However, it cannot be simply *any* effort or achievement that is ethically material in this situation. If a student is maintaining paid employment, this shows both effort and achievement, yet this does not seem an appropriate basis for assessing whether to offer unconditional admittance to a course of study. Instead, it must be effort and achievement where we can see a causal connection to what is being assessed. Examining students or looking at their recent academic work, for example, does better at capturing relevant effort and achievement than the use of characteristics that do not testify to this, such as ethnicity, age or gender. The question is, how can we capture this valuing of individual agency whilst employing learning analytics?

As noted in chapter 4, Kasper Lippert-Rasmussen (2011) claims that generalisation from groups is somewhere in the background of our assessments of people; for example, when we use an examination that has been normed on others' outcomes. Lippert-Rasmussen argues that given this, people cannot have a fundamental right to be treated *fully* as individuals, where this means judging people without reference to *any* generalisations. He suggests instead that part of treating someone as an individual could involve using information that categorises them as part of a group, if we are *also* careful not to ignore other information that is particular to that individual. This would suggest that learning-analytical assessment of students using group risk statistics can treat students more as individuals the more it can incorporate further information particular to individuals.

Ideally, we want students to be able to be treated in a way that recognises their past and current efforts and demonstrated abilities and capabilities, over other information. We also want institutional decision-makers to see students' future actions as open to being decided by them, that is, to see students' futures as open to students choosing actions that may or may not be congruent with their past. For decision-makers in higher education institutions, the practical ethical question is thus whether there are ways for learning analytics data to include more information particular to individual students, where this concerns their

choices, efforts and achievements. If so, this could be a way to help embed recognition of students' individual agency in institutions' use of learning analytics. I will briefly explain three paths that could be especially promising for education designers to explore in this regard, before concluding with a note of caution regarding focusing on students as individuals.

5.5 Recommendations
i) Nature of the risk factors – choices, efforts and achievements

Can the factors used in learning analytical assessment of students include more focus on individual choices, efforts and achievements by students that may be correlated with risk? For a general example, suppose that students who complete an online study timetable prior to starting a course are at lower risk of dropping out of their courses. Education designers could check whether students have the option of completing a study timetable for their courses, and if not, build this into the design of the course. A more specific initiative would require consideration of student choice or effort factors correlated with success in particular types of courses. Suppose that success for students studying humanities courses correlates with accessing library resources more than it does in science courses (see e.g. Sclater, 2014b). Education designers working on humanities courses can thus build in learning activities that require students to effectively access library resources across these courses. Including activities of this sort would help incorporate risk information that draws more on the choices and effort of individual students, in contrast to risk statistics for factors such as ethnicity or nationality.

ii) Nature of the risk factors – static or dynamic

The second question on the nature of the risk factors is whether the risk factors are *static* or *dynamic*. Static risk factors are largely historical, they do not change, or change little and only slowly (such as aging) (Witt, Dattilio & Bradford, 2011). Examples of static risk factors include factors from childhood, ethnicity or gender. Dynamic risk factors can change relatively quickly, such as a student's current courses of study or health status or internet search skills. Both static and dynamic risk factors can include factors for which a person is responsible and factors for which they are not. What is ethically salient here is that static

144

risk factors are not usually easily amenable to change. A student assessed as high risk on static risk factors, if assessed six months later on the same factors, will be still classed as high risk, regardless of any otherwise risk-reducing efforts they have made over this time. Static factors tend to be quicker and cheaper to assess, dynamic factors can be more expensive. But assessing only or primarily static factors may bypass possibilities for current and future agency on the part of the individual applicant, either on their own, or supported by others. Education designers working at a program level can check whether there are avenues for students to signal changes to dynamic risk factors, and how these may change the prompts for teachers (or courseware) to introduce academic intervention.

iii) Use of statistics specific to individuals

Schauer (2003) depicts institutional decision-makers as having two avenues available for assessing a person's risk, namely, expert examination of an individual or evidence from group risk statistics. There is a third avenue, however. Individual behavioural analytics uses a collection of behavioural data across time from the same individual. An example of this was given in chapter four, namely, the practice of 'experience rating' by insurers. This is where individuals' insurance premiums change depending on their behaviour in the previous year, rather than just on the risk associated for groups of which they are a member. Education institutions can capture data on individual student behaviour that can be correlated with outcomes for that student. We see this, for example, in the computer based Cognitive Tutor program that tracks individual student answers and responds with tailored content and activities (Mayer-Schönberger & Cukier, 2014), or in higher education institutions that offer online learning that tracks learners' behaviours at a more general level (see, e.g., Sclater, 2014b; Wagner & Hartmann, 2013).

Obviously, educational design is crucial to the performance of individual content responsive programs such as Cognitive Tutor, and this is a key area where educational design can support institutions to respond to students more as individuals. At the more general level, education designers can check whether online course pages can measure and display to students the amount of time a student has spent actively engaged on the page, and correlate this with the student's test score for items relating to the topic covered on the

page. This could provide a useful basis for students to reflect on their learning and study behaviours.

Comparing the three methods for assessing individual students to predict their future performance, we see that one involves experts examining individual students, the second involves assessment using statistical evidence derived from *extending other students' behaviour* over an individual student, and the third is assessment based on individual analytics, using statistical evidence derived from extending *the individual student's past behaviour* over that student. This third approach is a use of evidence that is both statistical yet specific to the individual student (see Colyvan, Regan & Ferson, 2001). I suggest that the more that individual behavioural analytics can be incorporated in risk assessment of students, the more learning-analytical risk judgments should be seen as recognising and including students as individual agents.

Institutions of higher education have the power to decide what to measure in terms of risk features for the problem of student dropout and failure. In this chapter I have focused on data where risks are imputed to students; however, this should not be taken to suggest that the risk of student dropout or failure is generated solely by students. Institutional decision-makers should be aware that how institutions choose to structure themselves also contributes to the risk of students not completing. For example, Kovacic (2010) found that two of the three biggest factors predictively distinguishing successful from unsuccessful students at a distance education institution were the *program* students were enrolled in and the *trimester* in which they were enrolled. Institutions can choose the extent to which they restrict their measuring of risk to the risks attributable to students, or instead include also statistics relating to their own programs, courses, schools or departments, and general processes and structure. This choice will influence who will bear the burden of addressing the risks: students or the institution.

The burdens involved for institutions to address institutional risk factors associated with increased student dropout or failure are the same as those for any organisational change. These include, for example, the financial costs of reorganising institutional structures, and staff feeling stressed or unhappy about the change. The possible burdens for 'risky' students

could include extra work or financial cost from requirements to take bridging courses, or restrictions on the number of courses, level of courses or pathways through a course. Ideally, from the point of view of increasing student success, we might hope institutions would focus on multifaceted interventions addressing both student and institutional risk factors. However, institutions will no doubt have a significant focus on the costs to themselves of any potential interventions. Richardson, Abraham and Bond (2012) suggest that cognitive changes in students such as increased effort regulation and elevated grade goals may be more cost-effective than multifaceted interventions, even if the latter might be academically better.

Higher education institutions are in a position of power relative to students, so it is possible that the institutions will mostly focus on students as 'risk bearers'. This would mean students would bear most of the burden of addressing risk. Given this concern, the transparency of institutions about their use of learning analytics is a significant issue. This has been recognised by higher education institutions; the Open University, for example, has developed plain English documents with a good level of detail that clarify how they will collect and use student information in their learning analytics (see Open University, 2014). This level of transparency is certainly not universal, however. Deploying learning analytics to risk assess students without transparency regarding the scope of its use maintains the power differentials between the institution and the students. It may mask institutional partiality in the way knowledge of the issue is constructed, and forestall institutional improvement.

Conclusion

In the education literature, current discussion on the ethics of the use of learning analytics has focussed particularly on principles drawn from existing ethical guidelines on institutional use of data. The discussion covers many of the ethical concerns learning analytics raises, especially the issues of transparency, consent, choice, accountability, privacy and security of data. This discussion has produced some excellent recommendations for ethical practice, including detailed policy recommendations (see, e.g., Open University, 2014). One ethical concern that has been less well covered is that of treating individuals differently on the basis

of group categorisation, particularly where individuals are burdened due to group based risk statistics. Schauer (2003) argues there is no significant difference between judgments based on 'individualised' evidence, and judgments based on statistics on group risk. In this chapter I have explored and challenged this argument. This critical exploration has resulted in the identification of distinctive ethical concerns regarding individuality and agency. As higher education institutions have good reason to use learning analytics to statistically assess risk regarding students' learning, I propose educational design approaches that may mitigate these ethical concerns. These are to consider the nature of the factors used in analytics and, where possible, to incorporate more use of factors involving individual effort and dynamic rather than static factors, and to make greater use of statistics specific to individual students. Finally, I support the importance of transparency by institutions regarding their use of analytics and measuring of what contributes to students' risk of dropping out or failing.

Chapter Five References

Alexander, Lawrence (1992) What makes wrongful discrimination wrong? Biases, preferences, stereotypes and proxies. *University of Pennsylvania Law Review, 141* (1): 149-219.

Arneson, Richard (2006) What is wrongful discrimination? *San Diego Law Review*, 43: 775–808.

Baker, Ryan & Siemens, George (2014) Educational data mining and learning analytics. In K.R Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*, (2nd ed., Pp.253-272). New York: Cambridge University Press.

Bichsel, Jacqueline (2012) Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations (Research Report). Louisville, CO: EDUCAUSE Center for Applied Research, August 2012. Retrieved from http://www.educause.edu/ecar.

Campbell, John, deBlois Peter, and Oblinger, Diana (2007) Academic analytics: A new tool for a new era. *EDUCAUSE Review, 42*(4): 41–57.

Colyvan, Mark, Regan, Helen and Ferson, Scott (2001) Is it a crime to belong to a reference class? *Journal of Political Philosophy*, *9* (2): 168-181.

Crawford, Kate, & Schultz, Jason (2013) Big data and due process: Toward a framework to redress predictive privacy harms. NYU School of Law. Public Law Research Paper No. 13-64: 93-128.

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York: Routledge.

Eidelson, Benjamin (2013) Treating people as individuals. In Deborah Hellman & Sophia Moreau (Eds.) *The Philosophical Foundations of Discrimination Law* (Pp.354-395). Oxford: Oxford University Press.

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, MA: Harvard University
 Press.

Information Commissioner's Office (n.d.) Data protection principles. Retrieved from
 https://ico.org.uk/for-organisations/guide-to-data-protection/data-protection-
 principles/

International Working Group on Data Protection in Telecommunications (IWGDPT) (2014)
 Working paper on big data and privacy: Privacy principles under pressure in the age of
 big data analytics. Retrived June 26, 2015 from
 http://dzlp.mk/sites/default/files/u972/WP_Big_Data_final_clean_675.48.12%20%28
 1%29.pdf.

Jia, Pengfei (2014) Using predictive risk modeling to identify students at high risk of paper
 non-completion and program non-retention at university. MBus thesis. Auckland
 University of Technology.

Johnson, Jeffrey A (2014) The Ethics of Big Data in Higher Education. *International Review of
 Information Ethics*. Retrieved from http://www.i-r-i-e.net/inhalt/021/IRIE-021-
 Johnson.pdf

Kay, David, Korn, Naomi and Oppenheim, Charles (2012). Legal, Risk and Ethical Aspects of
 Analytics in Higher Education. *CETIS Analytics Series*, 1 (6): 1-30.

Kovacic, Zlatko (2010) Early Prediction of Student Success: Mining Students Enrolment Data.
 Proceedings of Informing Science & IT Education Conference (InSITE) 2010.

Lippert-Rasmussen, Kasper (2006) Private discrimination: A prioritarian, desert-
 accommodating account. *San Diego Law Review, 43*: 817-856.

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15: 47-59.

Lokken, Fred, & Mullins, C (2015) ITC 2014 Distance Education Survey Results. Retrieved November 2015 from http://www.itcnetwork.org/membership/itc-distance-education-survey-results.html

Long, Phillip, & Siemens, George (2011) Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5): 31–40.

Mayer-Schönberger, Viktor and Cukier, Kenneth (2014) *Learning with big data*. Boston/New York: Houghton Mifflin Harcourt.

Moreau, Sophia (2010) Discrimination as Negligence, *Canadian Journal of Philosophy, 40:sup1*, 123-149.

Moreau, Sophia (2013) In Defense of a Liberty-based Account of Discrimination. In Deborah Hellman & Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law* (Pp.71-86). Oxford: Oxford University Press.

OAAI, Marist College (2012) Open Academic Analytics Initiative. Retrieved November 2015 from https://confluence.sakaiproject.org/pages/viewpage.action?pageId=75671025

Oblinger, Diana (2012) Let's Talk … Analytics. *EDUCAUSE Review*, 47(4): 10–13.

Open University (Oct 2014) Ethics use of student data for learning analytics policy FAQs. Retrieved October 2015 from http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-student-data-faq.pdf

Open University (Sep 2014) Policy on ethical use of student data for learning analytics. Retrieved October 2015 from http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-use-of-student-data-policy.pdf

Pardo, Abelardo and Siemens, George (2014) Ethical and privacy principles for learning analytics. *British Journal of Educational Technology, 45*: 438-450.

Prinsloo, Paul and Slade, Sharon (2014) Educational triage in open distance learning: Walking a moral tightrope. *The International Review of Research in Open and Distance Learning*. Retrieved from http://oro.open.ac.uk/40903/1/3007

Richards, Neil and King, Jonathan (2014) Big data ethics. *Wake Forest Law Review*, 49: 393–432.

Richardson, Michelle, Abraham, Charles, and Bond, Rod (2012) Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin*, 138(2): 353–387. DOI: 10.1037/a0026838

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Scholes, Vanessa (2016) The ethics of using learning analytics to categorize students on risk. *Educational Technology Research & Development.* DOI: 10.1007/s11423-016-9458-1

Sclater, Niall (2014a) Code of practice for learning analytics: A literature review of the ethical and legal issues. Jisc, London.

Sclater, Niall (2014b) Learning analytics: The current state of play in UK higher and further education. JISC. Retrieved from http://repository.jisc.ac.uk/5657/1/Learning_analytics_report.pdf

Sclater, Niall and Bailey, Paul (2015) Code of practice for learning analytics. Jisc, London. Retrieved October 2015 from https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics

Simpson, Ormond (2009) Open to people, open with people: Ethical issues in open learning. In Ugur Demiray and Ramesh C. Sharma (Ed.s) *Ethical Practices and Implications in Distance Learning* (pp.199-215). New York, USA: Information Science Reference.

Slade, Sharon and Prinsloo, Paul (2013) Learning Analytics: ethical issues and dilemmas. *American Behavioural Scientist*, 57 (10): 1509-1528.

Vuong, Annalies, Nixon, Tristan, & Towle, Brendan (2011) A method for finding prerequisites in a curriculum. From M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero & J Stamper (Eds.) Proceedings of the 4th International Conference on Educational Data Mining. Retrieved from http://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm2011_paper8_short_Vuong.pdf

Wagner, Ellen and Hartman, Joel (2013) Welcome to the Era of Big Data and Predictive Analytics in Higher Education. SHEEO (State Higher Education Executive Officers Association Higher Education Policy Conference 2013 [presentation] Retrieved from: http://www.sheeo.org/sites/default/files/0808-1430-plen.pdf

Willis, James and Pistilli, Matthew (April 7, 2014) Ethical discourse: Guiding the future of learning analytics. *EDUCAUSE Review.* Retrieved October 2015 from http://www.educause.edu/ero/article/ethical-discourse-guiding-future-learning-analytics

Willis, James (August 25, 2014) Learning analytics and ethics: A framework beyond utilitarianism. *EDUCAUSE Review*. Retrieved October 2015 from http://er.educause.edu/articles/2014/8/learning-analytics-and-ethics-a-framework-beyond-utilitarianism

Witt, Philip, Dattilio, Frank and Bradford, John (2011) Sex Offender evaluations. In E. Drogin, F. M. Dattilio, R.L. Sadoff & T. G. Gutheil (Ed.s) *Handbook of Forensic Assessment: Psychological and Psychiatric Perspectives*. Wiley Online.

Woodley, Alan and Simpson, Ormond (2014) Student Dropout: The elephant in the room. In Olaf Zawacki-Richter & Terry Anderson (Ed.s) *Online Distance Education: Towards a research agenda*. (Pp.459-483). Edmonton, Canada: AU Press.

Chapter six   Statistical discrimination in hiring[47]

Introduction

This chapter examines the differences between statistical discrimination and other ways of using information to make choices in hiring. Kasper Lippert-Rasmussen suggests hiring as an area where a moral case can be made in favour of statistical discrimination, if it proves an efficient way of hiring the best qualified applicant (Lippert-Rasmussen, 2007, p.386). Assessing job applications with the aim of improving the productivity of an organisation will involve attempting to predict the likely performance of applicants. A focus on efficiency in forecasting performance promotes group-level discrimination if there is group information that is statistically linked to productivity. What are the ethical concerns with employers aiming to identify and exclude applicants on group features that are statistically correlated with lower performance? In this chapter I draw conceptual and ethical analyses from consideration of the practical nature of statistical discrimination in hiring, with reference to the perspectives of both applicants and employers. I show how the key ethical issue of risk-based statistical discrimination can be viewed a conflict of liberties between applicants and employers: the liberty to access employment opportunities as an individual agent *versus* the liberty to access information to try to improve your workforce's productivity.

There is a significant body of research discussing how to hire to improve productivity (see, e.g., Le, Oh, Shaffer and Schmidt, 2007; Sackett and Lievens, 2008; Schmidt and Hunter, 1998), including identification of features of employees that are associated with higher and lower job performance. Statistical discrimination in hiring occurs when an employer takes a feature correlated with higher or lower employee productivity across the group that has that feature and uses it to assess individual applicants; for example, to discriminate against applicants who belong to a group with increased risk of lower performance. The group characteristics that could be investigated for links to productivity tend to fall into two categories: private activities and personal features. *Private activities* could include, for example, smoking or taking other drugs (legal or illegal), caring for children or elderly parents, religious activities, sports activities, and other leisure activities. *Personal features* could include, for example, race, gender, age, personality traits, other aspects of physical

---

[47] Sections 6.1-6.3 of this chapter contain some paragraphs that draw on Scholes (2014).

appearance. Some characteristics may cross the two categories: a person's weight might be partly due to genetics, whilst also influenced by private activities such as eating and exercise (Roehling, 2002, p.183).

Frederick Schauer (2003, pp39-41) notes that our moral judgements of applying group generalisations to assess individuals may be skewed by our awareness of a social history of discrimination against certain groups. This history has resulted in human rights legislation that prohibits employment discrimination on certain grounds. These grounds usually include sex, race, religious beliefs, ethnicity and national origin (Gandy, 2010, p.35), amongst others. It does not matter if research into productivity provides a rationale for statistically discriminating on any of these group features, this will be legally prohibited. Schauer follows a strategy of using an issue that avoids setting off our historically-aware moral reactions, in order to better focus on the conceptual issue of applying generalisations to judge individuals. To avoid illegitimately harnessing these moral reactions, most of the examples of statistical generalisations I use will avoid features commonly prohibited under human rights legislation. Lawful statistical discrimination in hiring could include screening for features such as smoking, residential demographics (good / bad neighborhood) (see, e.g., Nunn et al. 2010; Truth and Justice Commission 2011, 225; Cass and Garde 1984), and height (see, e.g., Schick and Steckel 2010), for example.

Whether a group characteristic correlates with productivity is a matter to be decided by empirical research. Currently, if empirical research on a characteristic either has not been done or has produced contradictory results, employers are faced with an uncertainty when it comes to making decisions using the characteristic. But once this research can assign a statistical probability to the relation between the characteristic and productivity at the group level, this creates a rational basis for statistical discrimination. This will effectively turn the uncertainty into a situation of quantifiable risk – thus creating a measurable incentive for employers to engage in statistical discrimination. So it is to be expected that the very pursuit of research into characteristics correlated with productivity will prompt increasing use of statistical discrimination in the future (Gandy, 2010).

After reviewing the literature on statistical discrimination in hiring, I will compare how statistical discrimination categorizes applicants relative to other hiring practices, to discern the practical nature of this discrimination in hiring. I will consider what statistical discrimination expresses about job applicants, and look at its impact on people's freedom to access a core life opportunity (cf Hellman, 2008, Moreau, 2010 and Eidelson, 2013, as outlined in chapters one and two). A couple of hiring scenarios will illustrate the implications for applicants, showing how risk-based statistical discrimination constrains applicants' autonomy and agency. This draws from analysis in chapter three. The final section of the paper focusses on the employer's perspective, and draws from analysis in chapter four. I show how the characterization of the employer's purpose in hiring, and the constraints on organisational decision-makers in a competitive business environment, influence the permissibility of risk-based statistical discrimination. I conclude that although risk-based statistical discrimination disrespects some applicants as autonomous agents, which is morally undesirable, it may be morally permissible for employers to engage in this discrimination. I end with some discussion of the extent to which recommendations for recognising and supporting individual agency (from chapters four and five) apply in the employment context.

6.1 Discussing statistical discrimination in employment

Discussion of statistical discrimination in employment has occurred predominantly in the economics literature. Seminal papers by Phelps (1972) and Arrow (1973) are credited with originating discussion of statistical discrimination; and Aigner and Cain (1977) also produced an early influential treatment (Fang and Moro, 2011, p.135). Typically theorists propose a scenario with applicant groups A and B. Members of Group A are rated by employers as having lower average productivity than Group B, either because Group A has lower average productivity, or because the employer finds it harder to estimate the productivity of Group A members. Using mathematical models that randomly assign applicants to employers, theorists found that the hypothesized statistical discrimination of a lower wage offer to Group A members would occur, and would persist over time. This finding was of significant interest. Previous discussion of employment discrimination in the economics literature focused on *taste-based* discrimination and argued that it was not supported by a

competitive market system so market forces would penalize employers for this and work toward eliminating it (Becker, 1957, 1971). With *statistical* discrimination, however, the unequal treatment and economic discrimination are actually generated by a competitive market system (Sattinger, 1998, p229).[48]

It is common in the economics literature to model the effect that statistical discrimination has on the *level of wage* an employer will offer an applicant. The economist's model of the hiring process includes an *employer* who needs a certain amount of labour and is willing to pay up to a certain price for it, and a series of applicants, each of whom wants to supply a fixed amount of labour beyond a certain price. The model seems to presume that applicants apply, and employers receive their applications, one at a time. The employer decides if the applicant in front of her is higher or lower quality and offers a wage accordingly, and so on for the next applicant, up to the point at which the amount of labour-for-money available is taken. However, this presumption is inaccurate if employers do not take in, review and decide on job applications on an individual basis, but instead have applicants present such that the employer can rank them *against each other* rather than just against a criterion for higher or lower quality. For example, the employer may advertise a position and set a deadline for applying before which decisions and offers will not be made. This expands the applicant pool for the employer to consider at a particular point in time. This should result in more high quality applicants being on offer, but also means it will take the employer more time to make decisions, as there are more applications to review. This is an incentive for employers to find a method to screen out applicants in a cost-effective way, without reducing overall quality in the applicant pool – a method statistical discrimination seems to offer.

---

[48] Toward the end of the twentieth century and in the first decade of the twenty-first century, economic theorists discussed whether such discrimination is an equilibrium that would persist over time (both theoretical discussion, LeBlanc, 1995; and empirical investigation, Altonji and Pierret, 2001, Feltovich and Papageorgiou, 2004). Sattinger (1998) considers whether the equilibrium from statistical discrimination counts as a discriminatory equilibrium for a group (rather than for an individual). Lundberg and Startz (1983) discuss whether it would reduce efficiency to have laws against statistical discrimination, while Schwab (1986) examines whether employer use of statistical discrimination actually does allocate resources more efficiently than if employers ignored the information. Albert and Cabrillo (2000) present a theoretical analysis of sex-based statistical discrimination in hiring, and compare policies that aim to prevent it. Blume (2006) includes consideration of whether policy moves such as hiring quotas and firm subsidies can be effective against widespread negative attitudes towards Group A members. Ewens, Tomlin and Wang (2011, p.30) claim to have shown the presence of statistical discrimination by landlords in the North American rental market, and suggest this could generalize to other situations, including job search.

Suppose that an employer is about to look through 23 applications for a job vacancy. Thirty persons submitted applications for the vacancy, but seven applications were screened out before they reached the employer. The employer had read some research suggesting that smokers are, on average, less productive, more costly employees than non-smokers (see e.g. Chadwick, 2006; Lecker, 2009; Robroek et al, 2011; Sherman & Lynch, 2013). A particular smoker, of course, might happen to be much more productive than your average non-smoker. Nonetheless, the employer has reason to believe that smokers, as a group, are less productive, more costly employees than non-smokers. The employer put in a filter to screen out the applications of any candidates – seven in this instance - who did not check the 'non-smoker' box on the application form when they submitted their applications. There has been much less discussion of employers engaging in statistical discrimination to make a binary decision to screen some applicants out of the hiring process (see, however, Harford, 2008). Yet this latter situation is both more typical, and more worrying if discrimination occurs (Sattinger, 1998, p.229). While it is bad to be offered a lower wage for a job due to statistical discrimination, it is worse to be excluded from further consideration for a job, and hence not get the offer of any wage level.

Economists have presented theoretical models that report the economic incentives for statistical discrimination (see, e.g. Klumpp & Su, 2013; Mialon & Yoo, 2014), but gathering empirical data on the impact of statistical discrimination is difficult (Moro & Norman 2003). The vast majority of field experiments on hiring discrimination are correspondences studies (Bertrand & Duflo, 2016, p6), where researchers send out fictitious resumes to employers that differ only on the variable being investigated, such as ethnicity, gender, unemployment spells or sexual orientation. The studies tend to find discrimination against minority groups, but determining whether this is statistical discrimination or taste discrimination is a harder task (Bertrand & Mullainathan, 2004; Harford 2008; Bertrand & Duflo, 2016). Nonetheless, recent research has shown evidence of statistical discrimination (see, e.g. Lahey, 2008; Kaas & Manger, 2010; Roberto & Greenberg, 2013; Masclet, Peterle & Larribeau, 2014; Rissing & Castilla, 2014; Yip & Wong, 2014; Fadlon, 2015). Kaas and Manger, for example, found discrimination against workers with Turkish-sounding names in the German labour market, and discovered the discrimination disappeared once reference letters were included that

161

testified to "soft" skills, such as capacity for teamwork, conscientiousness, affability (Kaas & Manger, 2010, p6). If employers harboured a distaste for workers from a Turkish background, this additional information should not make this sort of difference to employers' hiring practice.

Moving outside the economics literature[49], Amanda Baumle and Mark Fossett (2005) offer a conceptual analysis of statistical discrimination in employment. They contrast statistical discrimination with taste discrimination, consider some of the implications of statistical discrimination in employment, and suggest that policies used to combat taste discrimination are not likely to be effective against statistical discrimination. Baumle and Fossett are also interested in why statistical discrimination seems to be accepted in some circumstances but not in others. They note that the legal system's approach to statistical discrimination in the United States tends to differ depending on whether the discrimination occurs in the insurance context or the employment context. Baumle and Fossett (2005, p.1270) suggest that statistical discrimination may appear more acceptable the less it impinges on essential individual rights. If a society considers insurance less of a basic right for people than work, then statistical discrimination might be acceptable to that society in insurance but not in hiring. As evidence of this, Baumle and Fossett (2005, p.1270) note that in states in the US where car insurance is mandatory — so not having insurance impacts more fundamentally on people's lives, preventing them from being able to run a car — some states have prohibited the usual practice in the automobile industry of statistical discrimination based on sex.

In chapter four I compared statistical discrimination in a tenancy scenario with statistical discrimination in insurance, along three dimensions: individualisation, consequences for the applicant, and knowledge and expectations of the applicant. I think these dimensions apply similarly for statistical discrimination in employment. First, if the employer makes salient just one or two pieces of information to judge an applicant, the individuality of the applicant is very much ignored. As we saw in chapter four, this contrasts with insurers, who may take account of a great many pieces of information about an applicant – a far more individualized

---

[49] This paragraph and the next paragraph contain sentences drawn or adapted from Scholes (2014).

consideration. Second, as with tenancy seekers, there are no beneficial consequences for job applicants screened out due to statistical discrimination. In contrast, the actuarial knowledge of insurers can benefit rejected insurance applicants through signaling information to them on their risk; insurance applicants can then use this information to their benefit (for example, get their health checked).

Third, employment plays a fundamental role in the formation of a person's identity (Fadjukoff 2007, 32). An important focus (although it is to be hoped, not the sole focus) of schools and colleges when educating people and influencing their identity, is to develop competencies for workforce activities. In your earliest inquiries of strangers at a party you may encourage them to say something about their lines of work. When people apply for a job, it seems reasonable that it is the outcomes of their work-related activity that they put forward and expect to be judged on. They do not expect to be screened out of the job application process, or to have their application downgraded, due to statistics for a group of which they are a member, and regardless of their work-related activities. In this respect, the knowledge and expectations of job seekers differ from those of insurance applicants. Applicants for insurance are, in a sense, *choosing* the discriminatory practices in order to gain the insurance benefits, but applicants in the hiring context simply have the practice imposed on them. Job seekers are not usually aware that the personal information they provide may form the basis for some practices of statistical discrimination. The lack of awareness, and hence reduced freedom to consent or otherwise act on the matter, is one part of what makes statistical discrimination less acceptable in employment.

6.2 Types of statistical discrimination in hiring[50]

As noted in chapter one, there are several ways in which differences between groups can make it potentially beneficial for an organisational decision-maker to screen applicants using statistical discrimination. I will explain these in an employment context. A difference in the *average* productivity for one group compared with another group is one basis for discrimination. Suppose that Marama is expanding her data-entry business and needs to

---

[50] The first three paragraphs in this section are adapted from Scholes (2014).

hire staff. She knows there will be a large pool of applicants as the work is not highly skilled, and the job offers a comfortable environment with flexible working hours. Marama has read research suggesting the performance of the group of people who take some types of mind-altering drugs is more impaired than the performance of the group of non-users. She decides to require applicants to take a drug-test. Applicants who test positive or refuse the drug test will be screened out. (As a side-effect, Marama would also be discriminating against a couple of applicants who do not take drugs but who see the test as an invasion of privacy and refuse to take the test on these grounds; this may not bother her).

A second basis for statistical discrimination is where there is a higher *variance* in the productivity of members across one group compared with another. Imagine a city with a spread of suburbs, many of which have relatively homogenous sub-populations. Strattonside, however, has two distinct sub populations. Two-thirds of Strattonside comprises families of high socio-economic status whose adult members are generally *more* work-ready and supported for employment than populations in surrounding suburbs in the city. The remaining third of Strattonside, however, comprises gang families whose adult members are far *less* work-ready and supported for employment than the populations in surrounding suburbs. The job productivity average for Strattonside as a whole is no different from any other suburb; but there is a wider variance in the productivity of suburb members. *Ex hypothesi*, an employer who selects an applicant from Strattonside rather than another suburb has a higher chance of getting a higher productivity employee, but also a higher chance of getting a significantly lower productivity employee. Employers tend to be loss-averse (Dickinson & Oaxaca, 2009), and a loss-averse employer would have an incentive to choose a member from another suburb over a Strattonsider.

A third basis for statistical discrimination can be available even without any significant variability in the groups being compared. It involves variability in the results for groups due to the test or selection process (Albert and Cabrillo, 2000, p.8). This is where the test—or tester—for predicting performance produces more accurate predictions for one group than for another.[51] For instance, suppose that Murray gets a high volume of applications for a

---

[51] For a detailed discussion of this issue, see Schmidt and Hunter 1976.

couple of skilled jobs. He decides to do a quick initial sort of the resumes looking at the quality of the college or university the applicant attended. Murray is familiar with a reasonable number of colleges; he recognizes the names of familiar colleges as 'high quality' or 'low quality' colleges. But quite a few of the colleges mentioned by applicants are from other countries and are unfamiliar to him. Murray does his quick initial sort of resumes. When the recognized colleges are sorted, some go in the 'Yes' pile, and some in the 'No' pile (those recognized as 'high quality' or 'low quality' respectively). But for the 'unfamiliar' colleges, *all* go in the 'No' pile. Here the benefit to Murray from the discrimination is not due to any difference in average productivity from graduates of familiar versus unfamiliar colleges, but from the avoidance of the uncertainty surrounding unfamiliar colleges.

It should be noted that considerations of *efficiency* will control the potential practice of statistical discrimination in employment; in other words, whether the practice is an efficient means of producing the desired outcome. If a group characteristic is found to correlate with productivity, there is still the question of whether the cost to the employer of testing for it outweighs the benefits of knowing about it. An example of a low cost test is a question on an application form, such as "Do you smoke?" or "Do you have a criminal record?" A pre-employment drug test, by contrast, imposes a significantly higher cost. Bertrand, Chugh and Mullainathan (2005) suggest cognitive factors affecting decision makers can influence them to rely on their perceptions of a group characteristic. These factors might include being rushed, having multiple obligations competing for attention, and performing a task with a nonverbal automated response as its output - such as a manager considering resumes and putting them in a "Yes" pile or a "No" pile (Bertrand, Chugh & Mullainathan, 2005, p. 96).[52]

In conclusion, statistical discrimination in hiring can occur when an employer takes into account information about a characteristic that correlates with i) higher or lower productivity across the group of persons that has it; or ii) a high variance in the distribution of a productivity-correlated characteristic in one group; or iii) an uncertainty or pressure regarding measurement for a productivity-correlated characteristic for one group.

---

[52] Bertrand, Chugh and Mullainathan (2005) term this "implicit" discrimination rather than statistical discrimination. However, the results of their exploratory experiments to test for this, if extrapolated to real hiring situations, would fit the profile of statistical discrimination if making quicker decisions using the group characteristic was more profitable under these circumstances.

Employers thus use the group characteristic as a proxy for productivity information on applicants. The information is taken into account when considering employees from one group compared with those from other groups, in an attempt to reduce the risk of hiring a low productivity applicant.

6.3 Comparison with traditional hiring methods

Traditionally, common methods used to gain information from applicants in the hiring process include application forms, resumes, interviews and history and reference checks (Xuan Wang, 2011, p.53; Levashina & Campion, 2009, pp.236-239; Huffcut &Youngcourt, 2007, p.181; Taylor, Keelty & McDonnell, 2002, pp.9-10). For graduate selection in particular, assessment centres and work samples are also used (Bryson, James & Keep, 2013, p135). These methods typically produce information about qualifications; past work experiences and work ethic; skills and competencies; character (endorsements from referees); communication and social skills; and job applicants' stated ambitions and expectations of the position. Employers clearly believe that this constitutes relevant information about an applicant's job performance capabilities. How do traditional hiring methods differ from the other examples of statistical discrimination?

There may be a difference in the degree of *control* over the information provided. In most of the examples of traditional hiring methods, the applicant volunteers information of their choice to present themselves to the employer, such as their qualifications, past work experience and work-related activities or their chosen referees. In the letter accompanying their documents, the applicant can draw the employer's attention to the skills, experience and other attributes that speak to how they can contribute to the organisation. So the applicant has some control over how they present themselves to the employer. This contrasts with applicants' likely degree of control over presentation of themselves regarding information on a group feature requested by an employer for the purposes of statistical discrimination. For example, consider the difference between a section of an application form that asks applicants to outline their relevant work experience, with a section that asks applicants to check a YES / NO answer about smoking. The former allows the applicants

some control over how they present themselves to the decision-maker, whereas the latter does not. This may not be the case for all group features that an employer could use for statistical discrimination, however.

More importantly, there is a difference in the *locus of assessment*. Information about qualifications, work history and past or current work experience and skills is information about the individual's character and competencies that focuses on the *job-relevant performance* of the *individual applicant.* What the particular applicant has done and is doing are features of the individual that show that individual's character, skills and capabilities. Discrimination between applicants on these grounds is not statistical discrimination as the information is not a statistical proxy for productivity information about the individual, but instead a direct display of it. By contrast, with statistical discrimination, applicants are judged on a feature that relates to productivity at the *group* level, but has no necessary connection to the productivity of any particular individual applicant. The information that factor X is correlated with lower average productivity across the group with X (or that Xes) does not tell us of the particular performance of an individual applicant with X (or who Xes). This ignores the applicant's uniqueness. Moreover, if applicants are categorized on the basis of group features that do not involve the applicants' (past or present) chosen actions, this also ignores the applicants' autonomy.

Baumle and Fossett (2005) point out that decisions made on some of the traditional (legal) bases of discrimination, such as academic qualifications, could still involve statistical discrimination.[53] They give the example of an employer who screened out applicants for a job on the basis of whether the applicants had a college degree. Of course, screening job applicants based on college degrees does not necessarily involve statistical discrimination. Suppose that Employer B is looking to hire someone to fill a position as a lawyer, and that having a law degree is necessary to practice law in the position. If Employer B screens out applicants who do not have a law degree, this does not count as statistical discrimination against those applicants. By definition, no applicant without the qualification could be a

---

[53] Baumle and Fossett (2005, p.1255) note that discriminations based on group membership where membership is *achieved* rather than ascribed – for example, discrimination on the basis of 'having achieved a college degree' rather than 'being a particular race' – have not traditionally been seen as unfair discrimination.

good candidate; having the degree is an essential attribute for being able to do that job. Here the employer's discrimination is not based on a statistical correlation with productivity across a group, but rather a group feature that identifies a *minimum threshold* for being able to do the job.

However, suppose that Baumle and Fossett have in mind the following situation. Employer A has some vacancies for positions that do not require the sorts of skills that people can gain through completing an academic degree. Nevertheless, Employer A has seen research indicating that being a college graduate is statistically correlated with a higher level of productivity relative to the group of non-graduates. Employer A screens out applicants who do not have any college degree, based simply (or principally) on the belief that graduates, as a group, are more productive than non-graduates. If there are some applicants who are good candidates for the job but do not have a degree, these applicants will miss out on being considered for the job (Baumle and Fossett, 2005, p.1255). This would indeed count as statistical discrimination.

I have outlined two possibilities with regard to group discrimination concerning college degrees. The first was where a college degree is necessary for doing the job, and is not statistical discrimination. The second involved the rationale that college degrees are statistically correlated to increased productivity in a job, and is statistical discrimination. However, there is a third possibility. Suppose that Employer C has a vacancy for a job and is seeking applicants with a college degree. For this job, the degree qualification is not a legal requirement, but the *degree-type skills* that a person gains as part of achieving a college degree are necessary. This is probably the case for those jobs where employers may seek to hire graduates with degrees in humanities, for example. In this case, it is the degree-type skills that are job relevant, rather than the degree certification itself. If there are some applicants who possess the degree-type skills (and are thus good candidates), but do not have a degree, these applicants will not be considered by Employer C for the job.

The issue is how to determine whether an applicant has adequate degree-type skills, in order to be suitable for the job; and this involves the availability and reliability of information about the applicant. Obviously, a fast and reasonably reliable way of checking

for this particular information is for Employer C to look at whether someone has a degree. Presumably, an applicant who has the skills but not the degree could take steps to make information about their skills available to Employer C; for example, through obtaining a degree equivalency certification from a tertiary education institution through a process of Recognition of Prior Learning. The applicant needs to make an extra effort to notify Employer C of the relevant information about their skills, information that other applicants are able to provide simply through noting their degree. This is not unfair, as the other applicants had to make an extra effort to gain their degrees. For the generic higher thinking skills and writing skills such as can be developed through a college degree, the role and responsibility for credentialing these does not belong with an individual employer. As long as Employer C does not rule out in principle the consideration of alternative credentials that are shown to be equivalent, then Employer C would not be engaging in statistical discrimination in asking whether applicants had a college degree.

In conclusion, what distinguishes whether a particular hiring process counts as statistical discrimination is the relationship between *the information the process seeks,* and *what the job requires*. It is not statistical discrimination if the process seeks information that shows *that individual's* productivity with regard to the job requirements. It is also not statistical discrimination if the process identifies a minimum threshold for being able to perform the job, or seeks information that evidences skills needed in the job. It *is* statistical discrimination when a process seeks information about a factor that correlates with higher or lower productivity across a general group of people compared with another group and does not identify a minimum threshold required for performing the job. In the following two sections I will subject the two aspects of *the information the process seeks* and *what the job requires* to further ethical interrogation. These will be explored primarily from the viewpoints of the agents directly involved in the practices, namely applicants and employers, with particular focus on the issue of autonomous agency.

6.4 The information the process seeks
I will use two scenarios to illustrate potential impacts of statistical discrimination on applicants, showing privacy, liberty and agency concerns. Scenario One: Imagine an employer, Emma, who has read studies showing a statistical correlation between smoking

169

cigarettes and lower productivity in employees (see e.g. Chadwick, 2006; Lecker, 2009; Robroek et al, 2011; Sherman & Lynch, 2013). Emma is hiring in a weak labour market, with many applicants for each vacancy. Emma makes sure that whenever she advertises a vacancy, the application form asks whether the applicant smokes. When the forms come in, she screens out all the applicants who haven't answered "No" to the question about smoking. Suppose that Sam, a smoker applying for the job, realises such discrimination could occur. What could Sam do? Well, he could give up smoking, then truthfully answer "No" about smoking and remain in the running for the job. Or he could remain a smoker, and *lie* about it, again answering "No" to the question about smoking. Or Sam could remain a smoker and truthfully answer "Yes" to the question, or he could simply refuse to answer the question; either way getting struck off the list.

You might suggest there is another option for Sam in this scenario: Sam could simply drop the application for this job, and instead apply for another job elsewhere. However, recall that, like Emma, other employers have a rational incentive to engage in this discrimination. This means it is likely, once one employer starts doing this, that others will follow. This means that in exercising his liberty over providing his private information to employers, Sam is forced to choose between: 1) stopping his private, legal activity to maintain the same chance of getting a job; 2) lying about his private, legal activity to maintain the same chance of getting a job; 3) telling the truth and losing the chance to get this job and probably many others; or 4) refusing to answer and losing the chance to get this job and probably many others. Every option in this situation appears unpalatable. The first option imposes a serious restriction on Sam's liberty. The second requires Sam to tell an outright lie; this would be uncomfortable for him, and leaves him vulnerable in his job if the lie is later exposed. The third and fourth options both have the outcome that Sam will no longer be considered for this job, and may find it difficult to be considered for many other jobs.

This second scenario draws on the use of computer algorithms for predictive analytics. This is a relatively new practice for employment, although it has been used in fields such as marketing or financial services for longer; for example, data mining is used in financial services to identify clients with high-risk profiles (see, e.g. SAS Institute, n.d.). Scenario Two: Suppose that computer algorithms determine a strong correlation between higher levels of

productivity and a feature such as higher levels of 'online social networkedness'. Suppose too that large employers who have a dedicated Human Resources function start to ask applicants to supply a password for their private social networking pages, such as Facebook pages, in order to assess the amounts of 'online social networkedness' of applicants. The Human Resources team ignores the job applications from those applicants who refuse to comply with this request. This means job applicants who wish to protect the privacy of their personal lives - or applicants who, for other reasons, do not have private social networking pages - will be excluded from further consideration for the job vacancies.

You might object that this is unrealistic - it is not particularly obvious how higher levels of 'online social networkedness' could cause higher productivity. But this misses the point. For the employer, it does not matter if the causal connection is not obvious; as long as a strong statistical correlation holds, it is in the employer's interest to use it. As noted in chapter three, if applicants are categorized on the basis of a group feature with no obvious causal connection through the applicants to the organisational good, this stymies applicants' agency. Applicants (at least initially) will not expect to be assessed on the basis of 'online social networkedness', and certainly will not expect to be automatically rejected for having avoided Facebook or other social networking pages. We could even suppose that there is no actual causal connection between the two features – perhaps 'online social networkedness' is just a proxy for a third, unknown, feature that is causally connected to higher productivity. It does not matter. All that matters is that there is a strong enough correlation for the information to be useful to employers. If employers use a feature statistically correlated with productivity, but whose causal connection to job performance is either not obvious or non-existent, this bypasses the pathways through which applicants might try to develop themselves to be better at a job, and so stymies applicants' agency.

The thought of a world where job applicants can be rejected on the basis of a feature that lacks any discernible causal connection to applicant job performance is perhaps rather bewildering. We might think that if usable correlations can be determined, surely there is a causal connection that will be discernible with a bit more work? However, the algorithms that derive such correlations are not looking to accurately represent causal connections. Moreover, their predictions may not be reverse-engineerable to be scrutable to humans

(McKinlay, 2015). It is not the role of such algorithms to represent reality, but rather to identify patterns and make predictions from these, and how good they are is judged simply by how well they predict.[54] If a feature's connection to job performance were inscrutable, this would block applicants' agency in relation to developing their talents and capabilities for a job they desire to work at.

These two scenarios illustrate some privacy, liberty and agency concerns for applicants facing statistical discrimination. Note that these moral concerns are not exclusive to hiring situations involving statistical discrimination. Applicants asked to provide information directly relevant to job suitability might have privacy and liberty concerns too; for example, an applicant for a truck driving position who is asked to do an alcohol breath test might feel it is an invasion of privacy and that it imposes constraints on his liberty (to drink). The difference is in the connection to performance in the particular job, which is why the discrimination impacts on autonomy and agency. Applicants who cannot stay sober long enough to pass a breath test when applying for a truck driving job are not suitable to be driving trucks. They ought to know that, and can hardly claim the test as violating their autonomy. An applicant who smokes is unlikely to be unsuitable for a job simply through being a smoker[55]. Filtering out such applicants on a statistical correlation to productivity does constrain their autonomy. Moreover, as discussed in scenario two, if there is no discernible causal connection between a feature and job performance, there is no avenue for applicants trying to develop themselves on that feature in a way that makes them better at the job. If the only avenue for applicants is to try to develop the feature directly and hope to "cargo-cult" their way to a job, this blocks their agency.

6.5 What the job requires

The employer's liberty to treat job applicants in ways designed to increase workforce productivity raises fundamental questions about how we should conceive of a job. Organisations may see themselves as engaged in a private function when hiring, but society may consider the activity of seeking job applicants as at least in part a public activity. As well

---

[54] This is essentially the same point that Milton Friedman (1953) argued about economic models: it does not matter how (in)accurately the models represent the underlying reality, as long as they function to predict what we want predicted.

[55] With the exception perhaps of a job in a smoking cessation programme.

as providing developed social structures, institutions and infrastructure, societies allow business organisations particular benefits, such as a legal system with which to enforce business contracts, and limited liability in the event of failure. When hiring, an employer who publicly advertises a job attempts to benefit from having a group of appropriately-trained persons from which to source an employee. These applicants will have drawn on social resources for their education and training; hence the employer (amongst others) profiting from society's facilitation of their "human resources". Participation in a job is very important for most people's wellbeing and this participation is facilitated by society. If seeking employees is considered to some extent a public matter, then employers may be held to have responsibilities to undertake this with due regard to societal concerns, such as those raised in the previous section about risk-based statistical discrimination. My question is, given the implications of risk-based statistical discrimination for applicants, ought employers see themselves as fully at liberty (at least prima facie) to engage in this statistical discrimination? I will presume that if it is legal and we cannot reasonably characterise it as unfair, then employers can see themselves as at liberty to do this.

Alan Wertheimer (2006) notes the difficulty in distinguishing public and private spheres generally, as people may have different views based in their political ideas: the left tends to think sexuality and abortion to be private matters and discriminatory matters a public matter; the right tends to think the opposite (Wertheimer, 2006, p972). We could note that some organisations are seen as private, such as commercial, for-profit businesses, while other organisations are seen as public, such as government agencies, public hospitals, police stations, courts and state-funded universities. This seems to be the view of Peter Vallentyne (2006, p982), who contrasts *state* discrimination by state agencies or policies with *private* discrimination, which he defines as discrimination by individuals as citizens, including "discrimination in hiring, market exchange, and private association". Matt Zwolinski (2006), however, reserves the term *private* discrimination for personal association decisions such as whom to befriend or marry, and refers to discrimination in hiring and market exchange separately as *commercial* discrimination.

In support of Zwolinski's distinction, commercial businesses often differ from (other) individual citizens by having a public-facing presence in the community. Insofar as these

businesses have a public presence, the law tends to treat their actions as more akin to actions of public organisations, than to similar actions of private individuals. We see this, for example, in the distinctions made in employment discrimination laws. It may be objected that the law does not decide business responsibilities toward society. However, across the three predominant theories in the field of business ethics, namely Friedman's 'wealth-maximisation' theory, Freeman's 'stakeholder obligation' theory and Heath's 'market-failure avoidance' theory, compliance with relevant laws is the common minimum standard. Certainly, the law is a constraint on employers' liberties to discriminate. Human rights legislation covering discrimination treat organisational decision-makers differently from persons acting in their capacity as a private citizen. For example, compare the agent of a retail business hiring an employee to clean the workplace, with a private individual who hires someone to work in their own home as a cleaner. Employment law permits the private individual to discriminate on the basis of gender, religion, sexual orientation, political beliefs – they may discriminate in all manner of ways that the law prohibits to the agent acting for a business organisation.

While human rights legislation constrains the discriminations employers are permitted to make, it is sensitive to what business organisations need to know to successfully hire employees (Gandy, 2010, p.36). For example, while legislation may prohibit discrimination on the basis of religious belief and hence make it illegal for employers to ask questions about religious belief, exceptions are made for jobs where the religious belief is central to the work, such as being a vicar, priest or rabbi. Basically, employers are allowed to request information that is required for hiring purposes with regard to the job. I presume that if information is required in this way, employers cannot be characterised as unfair for using it. Hence, I will rework my question about employers' liberty to the following question: could the information about a factor that is statistically correlated with productivity be *required* for hiring purposes or hiring functions? To answer this, we will need to consider the characterisation of the employer's hiring activity and purpose, as well as the hiring environment.

What follows assumes employers operating in a Western capitalist economic environment. Levashina and Campion (2009, p.235) suggest it is a "fundamental truism in HR

management … that hiring procedures should be based on the job requirements", and that knowledge of job requirements ought usually to be obtained through a job analysis. Thompson and Smith (2009, p.260) suggest that "the purpose of hiring labor power is to expand capital". Together these suggest that the employer should monitor their workforce with an eye to the vacancy-filling or job creation that will expand capital, and set in place hiring procedures that reflect the requirements for the performance of those work positions. However, as it stands, this characterisation is too broad to help answer the question, as it is compatible with two different descriptions of the hiring purpose.

One description of the hiring purpose might simply be to source an employee who is suitable for the position. On the face of it, this description does not appear unreasonable. Employers do want an employee who is suitable for the work position, and employers are anxious to avoid hiring someone who is unsuitable. If we use this description of the hiring purpose, the question of whether the information about a factor that is statistically correlated with productivity is required for the hiring purpose can be answered with a clear "No". The standard hiring methods provide plenty of relevant information to identify a suitable applicant. While employers may nonetheless be interested in undertaking statistical discrimination as part of assessing applicants, the information could not be said to be *required* under this definition of the hiring purpose.

For vacancies that are difficult to fill, sourcing any suitable employee might be the extent of the hiring purpose. However, where there is a decent surplus of applicants employers may not want just *any* suitable employee. A second description of the hiring purpose is to hire the *best* – most productive - applicant for the position. If collecting some information to carry out statistical discrimination helps employers select a better employee than if it were not used, then employers could argue that it is necessary for the purpose or for carrying out the activity. This raises the question of whether employers should be considered to have a *right* to require information not strictly necessary for trying to find a suitable person for the job, but that could be necessary or useful for trying to find the *best* person for the job.

Rossler (2005, p.113) proposes that "If informational privacy consists in a person in principle being able to control or at least know or estimate who has what information about her, then

a violation of informational privacy consists ... in the person no longer having this control". Applicants clearly have an interest in having some control over the employer's access to at least some of their personal information. This is especially the case if employers might seek the information to use as a proxy for information statistically correlated with productivity at a group level. Applicants will want to ensure the information about them that is provided to the employer is accurate; that it is used to form an accurate picture of themselves; that only job-relevant information about them is used to make a determination of their suitability; and that the personal information required of them is kept to a minimum, to best protect their privacy in their lives outside work.

Employers could appeal to their general right to liberty to support a claim of a right to try to source the most suitable person for the job (rather than simply any person who could perform adequately in the position). The employer owns the business and is the direct cause of the availability of the job. Provided employers are not discriminating on the basis of grounds prohibited in human rights legislation, why should they not have a right to whatever information they deem necessary to finding the person likely to be best for the job? This may also produce better consequences. Besides employers feeling happier about having fewer restrictions, society may have an interest in the employer hiring the person who seems most likely to be the best applicant for the job. Having the best available people in jobs means an increase in productivity for businesses. Increased productivity may mean an expansion in business, creating more jobs and potentially increasing the taxes businesses pay to Government to spend on social issues.[56]

We have noted that human rights legislation that sets out prohibited grounds for discrimination places restrictions on what information an employer may legally use when making hiring decisions. This will only cover statistical discrimination that occurs on one of those prohibited grounds. If weight or smoking is classed as a disability or illness, such legislation will make it illegal to discriminate on these factors; but not otherwise. However, human rights legislation is not the only legislation covering the hiring process. In particular, privacy legislation sets out rules for the collection and use of personal information. This may

---

[56] This potential will not be realised if the 'best available people' include accountants who specialise in legal tax dodging.

bear on whether employers have a legal right to require applicants to supply any particular piece of information that employers might judge necessary for finding the best person for the job.

Privacy legislation will commonly take account of both the purpose for which information is collected, and the epistemic qualities of the information collected. For example, it may be specified that the information must be necessary for one or more of the activities or functions of the organisation (Australian Privacy Act 1988, S.1.1); and that with regard to the purpose for which it is used, the information is accurate, complete, up-to-date, relevant and not misleading (NZ Privacy Act 1993, S. II, p.8). Such stipulations make the employer's characterisation of their purpose central to assessing its legitimacy. Suppose that an employer states their purpose as 'to get a picture of the individual applicant X as a basis for making a hiring decision'. If the employer then statistically discriminates on several group features, this does not support their purpose. This is because statistical information that gives an overlay of a few group features produces at best a generic silhouette of a person, rather than anything approaching a picture of the individual applicant (a visual illustration of this is given in chapter 5). The group-level information also may not be accurate or relevant and so the picture may be misleading. However, suppose that the employer's purpose were restated as: 'to gain risk information on applicant X as a basis for making a hiring decision'. As compiling statistics on factors correlated with risk elements is exactly how risk profiles are built, the employer could claim that the information is fit for purpose and is accurate, relevant and not misleading.

Moreover, if the practice of statistical discrimination on a feature were to become widespread, then any employer might argue that they needed to engage in such discrimination to avoid being at a competitive disadvantage. For example, if the large companies in a particular industry adopted pre-employment drug-testing, small companies without such testing would have reason to fear that the applicants for their own vacancies would include a higher proportion of workers who would fail (or would perhaps have failed) these drug-tests. If this were the case, we might have to qualify our description of the employer's purpose. The purpose could be described as 'finding an employee who is no more likely to be a risky bet than the employees of one's competitors'. The purpose could

177

include more detail, such as 'engaging in a hiring practice that is no less risky than one's competitors in terms of both the employee sourced and the cost of sourcing'. Neither of these aims looks objectionable on the face of it. If these competitive conditions were to eventuate, an employer would have strong grounds for claiming that statistical discrimination was necessary for his or her purpose.

6.6 Recommendations

Drawing on analyses from chapter three, in chapter five I made some recommendations concerning the use of learning analytics to risk-screen students for potential for dropping out or failing. It may be instructive to consider those recommendations in the employment context. The recommendations were to consider the nature of the factors used in discrimination and, where possible, to incorporate more use of factors involving individual effort and achievement, plus dynamic rather than static risk-factors, and to make greater use of statistics specific to individual applicants. It seems the first recommendation on effort and achievement would be covered by focusing on the sorts of things employers have traditionally considered in hiring: applicants' qualifications and experience and demonstrated skills. The second recommendation was to focus on dynamic rather than static risk factors. This would support, for example, employers including time limits on tick boxes that ask about historical risk factors. An employer who has a blanket tick box question on whether an applicant has a criminal record could change the question to whether an applicant has had a felony conviction or been incarcerated in the last year.

The third recommendation was to make more use of statistics specific to individuals. The higher education context has two features that facilitate this. First, many individuals entering this context have statistics on their education history that may be reasonably relevant to this context. Second, individuals are regularly applying for admission to courses throughout their time with the education institution. Students apply at least every year, if not two or more times per year. This gives the opportunity to build a set of statistics specific to individuals from which individual risk factors may be able to be identified. It is harder to imagine the third recommendation working in this way in most employment contexts. Past employers may not want to release a history of detailed assessments they have made of

previous employees, preferring instead to give the briefer account that a reference provides. If the reference incorporates main points from more detailed assessments, that would help. However, it is likely that work assessments score the employee against a standard for performance. It is less likely that work assessments monitor the employee's strengths and weaknesses over time to facilitate identification of what features of the environment seem to have aided or hindered the employee. In this case, the assessments are less comparable.

In chapter five I also note the importance of institutions being transparent about their use of analytics and measuring of what contributes to students' risk of dropping out or failing. This is important in the employment context for similar reasons. Primarily, transparency is necessary for minimal consent to the process on the part of applicants. Also, employers are in a position of power relative to job applicants, and it may be in employers' interests that applicants bear most of the burden of addressing employment productivity risks. If the organisational structure or culture impacts on the productivity of workers, these elements may be hard or costly to try to change. What incentive has an employer to examine how the organisation supports success for applicants from 'riskier' groups? From an employer's point of view, it is easier to screen applicants than to assess organisational risk factors and institute organisational changes. Suppose that organisations implement statistical discrimination to risk assess job applicants without transparency about the scope of its use. Not only are applicants unable to consent, but this maintains the power differentials between employers and applicants. It may mask organisational bias in the way knowledge of employment productivity risk is constructed. A lack of transparency may hold society back from developing measures aimed at improving organisations to facilitate the employment of otherwise 'riskier' applicants.

Conclusion

From the viewpoints of the agents involved, the main ethical issue raised by statistical discrimination in hiring can be characterised as a conflict of liberties: the liberty to access information to try to improve your workforce's productivity *versus* the liberty to access

employment opportunities as an individual agent. Employers can require personal information from, and use it to discriminate against, a job applicant solely on the basis of a statistical link between such information and decreased productivity. This disrespects key features of applicants as agents, including their individuality, their consent and their autonomy. Against this, for employers to treat persons as individual agents at all stages of the hiring process may impose unreasonable costs; there may be competitive disadvantage to ignoring risk information; and the employer's goal of selecting the person likely to be most productive in the job has social benefits as well as self-interested benefits. When privacy legislation principles are based around the employer's purpose, hiring strategies aiming to select the best applicant appear to undercut these. Moreover, the competitive nature of the employment market means the use of statistical discrimination by some employers may put other employers in a situation where they feel they need to use it.

I conclude that, despite the disrespect with which excluded applicants are treated, under competitive labour market circumstances it seems morally permissible for employers to engage in statistical discrimination. Employers who are interested in trying to support applicants as individual agents can try to ensure their statistical assessments of applicants focus on features that show an individual's participation and achievements, plus tailor assessments to capture more dynamic factors, where possible. Interested organisations can self-evaluate on how their particular structures work to enable productive employment prospects for 'risker' applicant groups. Given the lack of a business rationale for this, perhaps employers or interested groups in society can ask industry groups to research this on their behalf. Finally, transparency is particularly important if employers use risk-focussed statistical discrimination. If employers are transparent about their use of this statistical discrimination, this will assist research into organisational structures and facilitate consent on the part of applicants.

Chapter Six References

Aigner, DJ and Cain, GC (1977) Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review,* 30: 175–187.

Albert, Rocio and Cabrillo, Francisco (2000) Sex discrimination: Is the labour market a market for lemons? Retrieved 15 September 2005 from http://www.ucm.es/info/iudem/cabrillo.DT00-4.pdf

Arrow, Kenneth J (1973) The Theory of Discrimination. In Ashenfelter, O and Rees, A. (Eds), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press.

Baumle Amanda K and Fossett Mark (2005) Statistical Discrimination in Employment: Its Practice, Conceptualization, and Implications for Public Policy. *The American Behavioral Scientist*. 49: 1250-1274.

Becker, Gary S (1957, 1971, 2nd ed.). *The Economics of Discrimination*. Chicago: University of Chicago Press.

Benn, Stanley I (1978) Human rights – for whom and for what? In E Kamenka and AE-S Tay (Eds) *Human Rights*. London: E Arnold.

Bertrand, Marianne, Chugh, Dolly and Mullainathan, Sendhil (2005) Implicit Discrimination. *American Economic Review*, 95: 94-98.

Bertrand, Marianne and Duflo, Esther (2016) Field experiments on discrimination. NBER Working Paper No. w22014, February 2016.

Bertrand, Marianne and Mullainathan, Sendhil (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94: 991-1013.

Blume, Lawrence E (2006) The Dynamics of Statistical Discrimination. *The Economic Journal*, 116: F480–F498.

Bryson, Jane, James, Susan and Keep, Ewart (2013) Recruitment and selection. In Stephen Bach and Martin Edwards (Ed.s) *Managing Human Resources: Human resource management in transition*, (Ch. 7, pp125-149), John Wiley & Sons Ltd.

Cass, B. and P. Garde (1984) Unemployment in the Western Region of Sydney: Job Seeking in a Local Labour Market, *SWRC Reports and Proceedings* 47: 3–43.

Chadwick, Karen L (2006) Is leisure-time smoking a valid employment consideration? *Albany Law Review*, 70:117-141.

Dickinson, David L and Oaxaca Ronald L (2009) Statistical discrimination in labor markets: An experimental analysis. *Southern Economic Journal, 76* (1): 16-31

Ewens M, Tomlin B and Wang LC (2011) Statistical discrimination or prejudice? A large sample field experiment. Retrieved from http://ewens.tepper.cmu.edu/wp-content/uploads/2011/04/statistical-discrimination2011.pdf

Fadlon, Yariv D (2015) Statistical discrimination and the implications of employer-employee race matches. *Journal of Labour Research, 36*: 232-248. DOI 10.1007/s12122-015-9203-2

Fang, Hanming and Moro, Andrea (2011) Theories of statistical discrimination and affirmative action: a survey. In Benhabib, J, Jackson MO and Bisin A. (Eds), *Handbook of Social Economics,* 1A: 133-200.

Feltovich, Nick and Papageorgiou, C (2004) An experimental study of statistical discrimination by employers. *Southern Economic Journal*, 70: 837-849.

Fernandez, Roberto M and Greenberg, Jason (2013), Race, Network Hiring, and Statistical Discrimination, in Steve Mcdonald (ed.) *Networks, Work and Inequality (Research in the Sociology of Work, Volume 24*), Emerald Group Publishing Limited, Pp.81-102.

Friedman, Milton (1953) The methodology of positive economics. In *Essays In Positive Economics*. Chicago: University of Chicago Press.

Gandy, Oscar H (2010) Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems. *Ethics and Information Technology,* 12 (1): 29-42.

Harford, Tim (2008) The dangers of rational racism. In *The Logic of Life,* London: Abacus

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, Mass., USA: Harvard University Press.

Huffcut, AI and Youngcourt, SS (2007) Employment Interviews. In DL Whetzel and GR Wheatson (Eds), *Applied measurement: Industrial psychology in human resource management*. Mahwah, NJ: Lawrence Erlbaum.

Kaas, Leo and Manger, Christian (2010) Ethnic discrimination in Germany's labour market: A field experiment. IZA Discussion Paper No. 4717, February 2010. Retrieved from http://ftp.iza.org/dp4741.pdf

Klumpp, Tilman and Su, Xuejuan (2013) Second-order Statistical Discrimination. *Journal of Public Economics*, 97 (8): 108-116.

Lahey, Joanna N (2008) Age, women, and hiring. *The Journal of Human Resources*, 43 (1): 30-56.

Le H, Oh I-S, Shaffer, J and Schmidt, F (2007) Implications of methodological advances for the practice of personnel selection: How practitioners benefit from meta-analysis. *Academy of Management Perspectives*, 21 (3): 6-15.

LeBlanc, Greg (1995) Discrimination in the labor market. *Canadian Journal of Economics*. 28: 702-17.

Lecker, Martin J (2009) The smoking penalty: Distributive justice or smokism? *Journal of Business ethics,* 84: 47-64.

Levashina, Julia and Campion, Michael A (2009) Expected practices in background checking: Review of the Human Resource Management Literature. *Employee Responsibilities and Rights Journal*, 21: 231-49.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *The Journal of Political Philosophy*, 15: 385-403.

Lundberg, S and Startz, R (1983) Private discrimination and social intervention in competitive labor markets. *American Economic Review*. 73: 340-7.

Maitzen, Stephen (1991) The Ethics of Statistical Discrimination. *Social Theory and Practice*, 17: 23-45.

Masclet, David, Peterle, Emmanuel, and Larribeau, Sophie (2012) The role of information in deterring discrimination: A new experimental evidence of statistical discrimination. Center for Research in Economics and Management, Working Paper 2012-28. Retrieved from http://crem.univ-rennes1.fr/wp/2012/201238.pdf

McKinlay, Steve (2015) The ethics of data convergence and predictive modelling. Paper presented at Australian Association of Professional and Applied Ethics Conference, 9-12 July, University of Auckland, New Zealand.

Mialon, Sue H and Yoo, Seung Han (2014) Employers' preference for discrimination. Emory Economics Working Paper Series 1402. Retrieved from http://economics.emory.edu/home/assets/workingpapers/smialon_14_02_paper.pdf

Moreau, Sophia (2010). Discrimination as Negligence, *Canadian Journal of Philosophy,* 40:sup1, 123-149.

Moro, Andrea and Norman, P (2003) Empirical Implications of Statistical Discrimination on the Returns to Measures of Skills. *Annales d'Economie et de Statistique*. 71-72: 399-417.

Nunn, Alex, Bickerstaff, Tim, Hogarth, Terence, Bosworth, Derek, Green, Anne E and Owen, David (2010), Postcode Selection? Employers' Use of Area- and Address-based Information Shortcuts in Recruitment Decision, Department for Work and Pensions Research Report 664.

Phelps, Edmund S (1972) The statistical theory of racism and sexism. *American Economic Review*, 62: 659-61.

Rissing, Den A and Castilla, Emilio J (2014) House of Green Cards: Statistical or preference-based inequality in the employment of foreign nationals. *American Sociological Review,* *79* (6): 1226-1255. Retrieved from http://asr.sagepub.com/content/79/6/1226

Robroek, Suzan JW, van den Berg, Tilja IJ, Plat, Jan F, Budorf, Alex (2011) The role of obesity and lifestyle behaviours in a productive workforce. *Occupational & Environmental Medicine, 68*: 134-139. DOI: 10.1136/oem.2010.055962

Roehling, MV (2002) Weight discrimination in the American workplace: Ethical issues and analysis. *Journal of Business Ethics*, 40: 177-89.

Rossler, Beate (2005) *The Value of Privacy*. Cambridge, UK: Polity Press

Sackett, Paul R and Lievens, Filip (2008) Personnel Selection. *Annual Review of Psychology,* 59: 419-50.

SAS Institute Inc, (n.d.) Who's using it? Machine learning: What it is and why it matters. [webpage] http://www.sas.com/en_us/insights/analytics/machine-learning.html#machine-learning-users

Sattinger, Michael (1998) Statistical Discrimination with Employment Criteria. *International Economic Review*, 39: 205-30.

Schick, Andreas and Steckel, Richard H (2010) Height as a Proxy for Cognitive and Non-cognitive Ability. *NBER Working Paper* No. 16570.

Schoeman, Ferdinand D (Ed). (1984). Privacy: Philosophical dimensions of the literature. In *Philosophical Dimensions of Privacy: An Anthology*. Cambridge: Cambridge University Press

Scholes, Vanessa (2014) You are not worth the risk: Lawful discrimination in hiring, *Rationality, Markets and Morals*, 5: 13-29.

Schmidt, Frank L and Hunter, John E (1998) The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124: 262-74.

Schwab, S (1986) Is Statistical Discrimination Efficient?. *The American Economic Review*. 76: 228-34.

Sherman, Bruce W and Lynch, Wendy, D (2013) The relationship between smoking and health care, workers' compensation and productivity costs for a large employer. *Journal of Occupational and Environmental Medicine, 55* (8): 879-884. DOI: 10.1097/JOM.0b013e31829f3129

Simpson, Ormond (2009) Open to people, open with people: Ethical issues in open learning. In Ugur Demiray and Ramesh C. Sharma (Ed.s) *Ethical Practices and Implications in Distance Learning* (pp.199-215). New York, USA: Information Science Reference.

Taylor, Paul, Keelty, Yvonne and McDonnell, Bridget (2002) Evolving Personnel Selection Practices in New Zealand Organisations and Recruitment Firms. *New Zealand Journal of Psychology*. 31: 8-18.

Thompson, Paul and Smith, Chris (2009) Waving, not drowning: Explaining and exploring the resilience of labor process theory. *Employee Responsibilities and Rights Journal*, 21: 253-62.

Truth and Justice Commission (2011) Volume One Report of the Truth and Justice Commission. URL: http://www-previous.usip.org/files/ROL/TJC_Vol1.pdf [29.07.2013].

Vallentyne, Peter (2006) Left libertarianism and private discrimination. *San Diego Law Review, 43*: 981-994

Wang, Xuan (2011). Characteristics of the human resource executives and companies that use more sophisticated employee selection methods. Thesis, Emporia State University. Retrieved 25 July 2016 from https://esirc.emporia.edu/bitstream/handle/123456789/998/XuanWang%20thesis.docx?sequence=1

Wertheimer, Alan (2006) Reflections on discrimination. *San Diego Law Review*, 43: 945-980

Yip, Chi Man and Wong, Raymond Sin-Kwok (2014) Gender-oriented statistical discrimination theory: Empirical evidence from the Hong Kong labor market. *Research in Social Stratification and Mobility*, 37: 43-59.

Zwolinski, Matt (2006). Why not regulate private discrimination? *San Diego Law Review*, 43: 1043-1061**.**

Chapter seven   Concluding comments

7.1 Legislation to address statistical discrimination

In this thesis, I kept my focus on the perspective of the two parties most directly affected by statistical discrimination, namely, the applicants who would be identified and treated differently because of it and the organisational agent who would be implementing the statistical discrimination. While analysing the moral undesirability of statistical discrimination in terms of how disadvantaged applicants are treated, I have tried to be sensitive to the interests and concerns of organisational agents, and to be pragmatic about the likely use of statistical discrimination. To this end, I included recommendations for how organisational agents could assess their use of statistical discrimination to better recognise and respect applicants as individual agents in the application process. However, I avoided taking a broader policy perspective of recommending changes to law or other government action. This is because I believe empirical investigation and local knowledge is required to understand how social policies will impact on undesirable statistical discrimination. I will briefly discuss one recent, prominent example to illustrate this, before concluding this thesis.

Criminal offenders have trouble integrating into society on their release from prison. Employment is one area where they suffer discrimination. Employers who include a tick box on their application form asking if applicants have a criminal record are not typically looking to practice affirmative action toward offenders. Employers' prejudice is not ill-founded: offenders, particularly if they have been recently incarcerated, are less productive, as a group, than applicants with no criminal history. They are less work-ready, in worse health and more likely to (re)offend (Bloom, 2006; Doleac & Hansen, 2016). Of course, some ex-convicts may make great employees given the chance[57]; but the application form screening out people with a criminal record means these 'suitable' applicants do not get the chance. Lacking the ability to gain employment in this way and the integration into society that employment facilitates, their environment becomes more conducive to re-offending (Henry & Jacobs, 2007). Employment is one of the biggest factors helping offenders desist from re-

---

[57] Many persons with minor criminal convictions do not get incarcerated for those convictions; these people may also get caught out by employer criminal record tick boxes.

offending (McSweeney & Hough, 2006). People interested in helping ex-prisoners rehabilitate have suggested legislation to ban employer use of the tick box on criminal record. This is known in the United States as 'Ban the box' policy (Henry & Jacobs, 2007).[58] The call to 'Ban the box' has also been supported by people working against racial discrimination in the United States. Hispanic and especially Black people are convicted and incarcerated at a far higher rate in the United States than other ethnic groups, and the rate is high for young Hispanic men and particularly high for young Black men. In some States, nearly one in three young Black men has a conviction and one in every four young Black men is or has been incarcerated. When these young men apply for jobs, if employers use the 'criminal record' tick box, the young men will find it extremely difficult to gain employment. Banning the tick box would thus seem not only to help ex-prisoners' chance of rehabilitation generally, but to help the job chances of Black men in particular. For these reasons, over 20 states and more than 100 jurisdictions in the United States have implemented 'Ban the box' legislation (Agan & Starr, 2016).

In two recent papers, researchers investigate the impact of Ban the box legislation in the United States, taking advantage of its different implementation timing across different states. Amanda Agan and Sonja Starr (2016) present the results of a field experiment in which they sent 15,000 fictitious resumes to employers in two jurisdictions (New York and New Jersey) pre- and post-implementation of Ban the box legislation. The resumes vary information on race (signalled through typically White or typically Black names) and a criminal record of a felony conviction[59]. During the pre Ban the Box (BTB) legislation period, the researchers find substantial discrimination against applicants with the criminal record. They find banning the box (BTB) increases the employment call-backs of applicants with criminal records – but only if they are white. A key finding was that "before BTB, white applicants to BTB-affected employers received 7% more callbacks than similar black applicants, but after BTB this gap grew to 45%" (Agan & Starr, 2016, p.4). This greater exclusion of Black applicants after the introduction of Ban the box legislation affects both

---

[58] This legislation does not prevent employers asking about criminal history later in the application process; it just prohibits having a tick box or question that removes applicants at the initial stage of the employment process without further consideration.

[59] The researchers also vary the CVs on type of educational qualification, and a one-year gap in employment history.

Black applicants with a criminal record and Black applicants who have a clean record. The researchers determine that the employers in their field experiment engaged in statistical discrimination in favour of white applicants and against black applicants. This indicates employers' discomfort with the removal of the information that the criminal record tick box offered.

Jennifer Doleac and Benjamin Hansen (2016) studied the effects of Ban the box legislation on the actual rates of employment for low skill jobs in jurisdictions, comparing the employment rates pre- and post- legislation, and exploiting the fact that different jurisdictions introduced the legislation at different times. The researchers found that Ban the box legislation decreased the employment chances of young low-skilled Hispanic men by nearly 3 percent, and decreased the employment chances of young low-skilled Black men by just over 5 percent. The researchers interpret their data similarly to Agan and Starr: Doleac and Hansen suggest that in the absence of the information provided by the criminal record tick box, employers statistically discriminate against all young male applicants who are Black or Hispanic. This means some young male applicants who are Black or Hispanic and who have a clean record do not get considered for a job. This is obviously not what was expected or wanted by the theorists who proposed the social policy changes resulting in this legislation.

Agan and Starr suggest a solution to the problem created by Banning the box could be to make the application process 'name-blind'. This means anonymising the applicant information so that employers who were inferring the race of an applicant from their name would no longer be able to engage in this racial discrimination. I would note that mandating an anonymising process imposes yet more legislation on employers. Employers may resist, especially if measures to ensure a blinded hiring process are more complicated and costly than removing a criminal record tick box from an application form. Discrimination on the basis of race is illegal, of course, whether the discrimination is statistical or taste based. However, it is not clear that most of the employers are engaging in illegal racial discrimination. Doleac and Hansen's analysis suggests that many of the jobs where young Black men lost out were likely to go to higher educated Black women or low skilled older

Black men, rather than just to White applicants. While still discrimination, it can hardly be classed particularly as *racial* discrimination if an employer discards the applications of young Black men in favour of older Black men, or Black women.

While social policy relating to statistical discrimination (and discrimination more broadly) requires empirical investigation into local conditions, I hope that the analysis I have provided in this thesis might be of some use. For example, the point above about employers' actions supports my argument in chapter two that discrimination can be about combining narrower stereotypes rather than it simply reflecting broader racial stigma. It seems it is the *combination* of being young and Black (or Hispanic) and male that is leading employers to worry about criminality, not race by itself. This raises the concern that if the application process for low-skilled jobs were name-blinded but not sex-blinded, perhaps employers would start avoiding young men generally. This is speculation. Nevertheless, perhaps this sort of speculation can point toward useful questions for empirical research in advance of implementing social policies to address undesirable statistical discrimination.

There is also the concern that employers who practice race-based affirmative action will be forestalled in their efforts if anonymisation is in place[60]. If such employers provide a welcoming and inclusive environment for Black workers it would be a pity to stymie this. An alternative to Agan and Starr's proposal can be found at the end of chapter 6 of my thesis. I suggested employers should make the information sought by criminal tick boxes more dynamic, by, for example, narrowing their question to whether applicants have been in prison in the last 12 months. This will not help applicants who have just been released from prison (at least, not straightaway)[61]. However, it should help some ex-prisoners, regardless of whether they are Black or White, without impacting on all the young Black men with clean records.

---

[60] Behaghel, Crepon and Barbanchon (2015) found that when French public service organisations could choose to participate in a process that anonymised the resumes of applicants, minorities had less chance of being hired than previously. Analysis from the researchers suggests that organisations that are inclined toward some kind of affirmative action may have disproportionately agreed to participate in the research. When these organisations received anonymised resumes, they could no longer give any form of preferential reading of the resumes of minority applicants, and so their hiring of minority applicants decreased.

[61] If support services can organise an option for ex-prisoners to be able to attend paid preparation-for-work programmes for 12 months upon release, this might help the recently released prisoners.

7.2 Conclusion

As we have seen, organisational decision-makers can use statistical discrimination to exclude or otherwise disadvantage applicants simply on the basis of their membership in a 'risky' group. With the exception of Kasper Lippert-Rasmussen (2007, 2011, 2014), recent philosophical literature on discrimination has tended to neglect statistical discrimination. In this thesis, I argued that risk-focussed statistical discrimination is a significant and morally problematic form of discrimination. This thesis helps theorists and decision-makers understand the drivers and ethical dimensions of this discrimination and offers recommendations for how to mitigate ethical concerns. The thesis does this through explanation, analysis and argument, as follows:

Chapter 1

Chapter one explained my choice of a pluralist, pragmatic methodology, and the background conditions that call for a defence of the moral significance of statistical discrimination. The distinctive moral concern with statistical discrimination is centrally bound up with identifying individuals as members of groups (namely, groups posing a higher risk to an organisation). I argued that disadvantageous exclusion on the basis of group generalisations is central to discrimination. I also argued that our concept of discrimination requires that it be *a potentially wrongful exercise of a legitimate practice*. The legitimate practices discussed in this thesis are insurance, admission to higher education, and hiring.

Chapter 2

Chapter two was an exploration and defence of the moral significance of statistical discrimination. I rejected conceptions of discrimination which restrict it to actions disadvantaging certain socially salient groups; this expands the ground covered by statistical discrimination, as group risk features need not align with recognition of a group as socially disadvantaged. I drew on empirical research to discredit accounts of discrimination that focus solely on the false beliefs or hostile attitudes of agents. I drew further on research on stigma and stereotypes from economics, sociology and psychology to show the force of stereotyping in discrimination. There are incentives for organisations to exclude or disadvantage persons through applying rational stereotypes on group risk in place of using

individualised evidence. As rational stereotyping underlies statistical discrimination, this further supported my argument for the significance of statistical discrimination.

Chapter 3

In chapter three, I turned from making a general case for the significance of statistical discrimination to arguing for the moral importance of this discrimination in organisational selection of applicants. I argued in favour of a sharper distinction between statistical evidence and individualised evidence than some other analyses suggest, and showed the implications for the autonomy and agency of applicants in organisational selection decisions. I argued contra Ferdinand Schoeman (1987) that decisions based in statistical probabilities are simply extended over an individual; there is a sense in which they are not really about the individual *qua* individual. I argued contra Frederick Schauer (2003) that because test-based evidence involves the participation of the individual and is designed to show the individual's capability (and perhaps also their character, depending on the test) it is different in kind from evidence comprised solely of statistics for group risk features. I also noted the asymmetric power that organisations have to locate risks in statistics relating to the attributes of applicants, rather than consider how the organisation's structures, policies, and cultures contribute to the risk they want to minimise.

I discussed Kasper Lippert-Rasmussen's (2011) argument that some use of 'individualised' evidence can simply reflect statistical discrimination operating at a deeper level, and his 'revisionist' version of what it is to treat someone as an individual that can include both statistical and individualised information. Drawing on Benjamin Eidelson (2013) and David Edmonds (2006) on what is valuable about treating a person as an individual, I presented my view that trying to treat people as individuals is morally important (at least in part) because it allows us to prioritise recognising a person as a locus of unique moral agency. This is forestalled if organisational decision-makers exclude or disadvantage people purely on the basis of group risk statistics, which is a reason to disprefer statistical discrimination. I suggested that if statistical evidence in these decisions displaces individualised evidence, decision-makers may not be able to treat an applicant as an individual, denying the applicant's unique autonomous agency.

Chapter 4

Being able to treat individual applicants as autonomous agents is morally important, but it is not the only morally important factor in organisational decisions. In chapter four, I proposed that the morality of statistical discrimination varies depending on several dimensions of the decision-making context. First, an organisation's purpose in society is ethically significant to evaluating its use of statistical discrimination to assess individuals. Second, the extent to which statistical discrimination is 'discriminatory' varies depending on how narrowly the assessment picks out the applicant. Third, the moral concern with risk-based statistical discrimination can vary depending on the provision of benefits and the transparency of the organisation's use of the method (as transparency affects whether applicants can give informed consent). Finally, I argued that the need for organisational decision-makers to use risk-focussed statistical discrimination changes depending on the information the decision-maker can access and the resources they have to spread risk.

In chapters five and six I discussed two contexts which provide evidence to support my argument that the decision-making context affects the morality of risk-focused statistical discrimination.

Chapter 5

I examined the use of learning analytics to assess the risk of tertiary education students dropping out or failing. This differs from risk-screening in more purely commercial contexts. In particular, the purpose of the organisation and benefit to applicants differs, which affects the morality of the practice. Comparing the treatment of an applicant under analytics risk-screening with a more traditional individualised assessment of a student, I found the learning analytics screening to be potentially ethically defective, lacking respect for students as individual agents. Nonetheless, as discussed previously, other dimensions of the decision context such as benefits to individuals affected and transparency can affect the morality of risk-based statistical discrimination. Learning analytics can bring important benefits for students and the organisation, and tertiary education organisations will use analytics, so I considered how they may be employed to better respect students as individual agents. My three recommendations were to focus on the extent to which the statistical data draws on factors of effort and achievement; uses dynamic rather than static factors; and draws from

the student's own history. These affect how well the organisational decision-maker respects the individual agency of students applying for admission to courses. I concluded with a note of caution on how the power asymmetry between the organisation and the student may affect how the risk of student drop out or failure is constructed, and the need for transparency.

Chapter 6

Chapter six presented the second of the two contexts supporting my argument that the decision-making context affects the morality of risk-focused statistical discrimination. I looked at statistical discrimination in employment. I compared this statistical discrimination with more traditional hiring practices, and analysed the key features of this discrimination in employment. I identified that statistical discrimination is when the hiring process seeks information about a productivity group risk factor that does not identify a minimum threshold required for performing the job. I then investigated the two dimensions of *the information the process seeks* and *what the job requires* from the viewpoints of applicants and employers, respectively. I suggested the main ethical issue raised by statistical discrimination in hiring can be characterised as a conflict of liberties: the liberty of employers to access information to try to improve their workforce's productivity *versus* the liberty of applicants to access employment opportunities as an individual agents.

Using a couple of scenarios, I illustrated the potential impacts of statistical discrimination on applicants. I showed the disrespect for applicants as agents, in terms of their individuality, their consent and their autonomy. I then discussed employers' purposes in hiring employees, and how the characterisation of this purpose affects whether risk-focussed statistical discrimination is necessary for employers. I noted some benefits to this discrimination, and how competitors' use of risk-focussed statistical discrimination may make it costly for employers to avoid using statistical discrimination. I concluded that it may be morally permissible for employers to engage in risk-based statistical discrimination, even if this has the morally undesirable effect of disrespecting applicants as autonomous individual agents. I finished with some recommendations for employers to mitigate this. These largely mirror the recommendations from chapter five, with particular mention of the importance of transparency.

7.3 Summation

My key question was: what distinctive ethical concerns inhere in the practice of organisations discriminating against applicants on the basis of group risk statistics? I argued that the distinctive concern is the lack of respect for applicants as unique autonomous agents. I argued further that the decision-making context affects the morality of risk-focussed statistical discrimination. Other things being equal, the morality of this discrimination varies depending on the purpose of the organisation, the level of detail in the discrimination, and whether the discrimination is transparent to applicants and includes some benefit for applicants. To mitigate the lack of respect for applicants as individual agents, organisational decision-makers can focus on the extent to which the statistical data comprises i) factors that feature effort and achievement; ii) dynamic rather than static factors; and iii) data drawn from the applicant's own history and actions over time.

Chapter Seven References

Agan, Amanda and Starr, Sonja (2016) Ban the Box, Criminal Records, and Statistical
Discrimination: A Field Experiment. Becker Friedman Institute for Research in
Economics Working Paper No 2016-17. Retrieved 1 November 2016 from
http://bfi.uchicago.edu/sites/default/files/research/2016-17.pdf

Behaghel, Luc, Crepon, Bruno and Le Barbanchon, Thomas (2015) Unintended effects of
anonymous resumes. *Americal Economic Journal: Applied Economics*. 7 (3): 1-27.

Bloom, Dan (2006) Employment-focused programs for ex-prisoners: What have we learned,
what are we learning and where should we go from here? MDRC Discussion paper, May
24, 2006. Retrieved from http://files.eric.ed.gov/fulltext/ED493008.pdf

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York:
Routledge.

Eidelson, Benjamin (2014) Treating people as individuals. In Deborah Hellman and Sophia
Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law*. (Pp354-395) Oxford
University Press: Oxford.

Doleac, Jennifer and Hansen, Benjamin (2016) Does "Ban the box" help or hurt low-skilled
workers? Statistical discrimination and employment outcomes when criminal histories
are hidden. NBER Working Paper 22469, July 2016.

Henry, Jessica and Jacobs, James (2007) Ban the box to promote ex-offender employment.
Criminology and Public Policy, 6 (4): 755-762. Retrieved 1 November 2016 from
http://www.nelp.org/content/uploads/2015/03/Henry-Jacobs.BantheBox.article.Oct-07.pdf

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *The
Journal of Political Philosophy*, 15: 385-403.

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15, 47-59.

Lippert-Rasmussen, Kasper (2014) Statistical discrimination. In *Born Free and Equal: A philosophical inquiry into the nature of discrimination*. Oxford Scholarship Online DOI:10.1093/acprof:oso/9780199796113.001.0001

McSweeney, Tim and Hough, Mike (2006) Supporting offenders with multiple needs: Lessons for the 'mixed economy' model of service provision. *Criminology and Criminal Justice*, 6: 107-125. DOI: 10.1177/1748895806060669

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Schoeman, Ferdinand (1987) Statistical vs direct evidence. *Nous*, 21 (2): 179-198.

References for all chapters

Abbring, Jaap, Chiappori, Pierre-Andre & Zavadil, Tibor (2008) Better safe than sorry? Ex ante and ex post moral hazard in dynamic insurance data. Tinbergen Institute Discussion Paper, TI 2008-075/3. Retrieved 28 October 2016 from http://dare.ubvu.vu.nl/bitstream/handle/1871/12965/08075.pdf?sequence=1

Agan, Amanda and Starr, Sonja (2016) Ban the Box, Criminal Records, and Statistical Discrimination: A Field Experiment. Becker Friedman Institute for Research in Economics Working Paper No 2016-17. Retrieved 1 November 2016 from http://bfi.uchicago.edu/sites/default/files/research/2016-17.pdf

Aigner, DJ and Cain, GC (1977) Statistical Theories of Discrimination in Labor Markets. *Industrial and Labor Relations Review*, 30, 175–187.

Albert, Rocio and Cabrillo, Francisco (2000) Sex discrimination: Is the labour market a market for lemons? Retrieved 15 September 2005 from http://www.ucm.es/info/iudem/cabrillo.DT00-4.pdf

Alexander, Lawrence (1992) What makes wrongful discrimination wrong? Biases, preferences, stereotypes and proxies. *University of Pennsylvania Law Review*, 141 (1), 149-219.

Allen, Ronald (2015) A note to my philosophical friends about expertise and legal systems. *Humana.Mente Journal of Philosophical Studies*, 28: 71-86.

American Academy of Actuaries (2006) The roles of the actuary in the selection and application of actuarial models: concepts on professionalism. Discussion paper no 7. Retrieved September 2106 from http://www.actuary.org/files/publications/models_june06.pdf

Arneson, Richard (2006) What is wrongful discrimination? *San Diego Law Review*, 43: 775–808.

Arrow, Kenneth J (1973) The Theory of Discrimination. In O Ashenfelter and A Rees (Eds), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press.

Arthur, Jeffrey B (1994) Effects of human resource systems on manufacturing performance and turnover. *Academy of Management Journal*, 37 (2): 670-687.

Avraham, Ronen, Logue, Kyle D and Schwarcz, Daniel (2013) Understanding insurance anti-discrimination laws. Public law and Legal theory Research Paper series, paper no. 289, Law & Economics Research Paper Series, paper no. 12-017: 1-52.

Baker, Ryan & Siemens, George (2014) Educational data mining and learning analytics. In K.R Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*, (2nd ed., pp.253-272). New York: Cambridge University Press.

Baker, Tom (2002) Risk, insurance, and the social construction of responsibility. In Tom Baker and Jonathan Simon (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (pp33-51). Chicago and London: University of Chicago Press.

Balliet, Daniel, Wu, Junhui & De Dreu, Carsten K W. (2014) Ingroup favouritism in cooperation: A meta-analysis. *Psychological Bulletin*. Doi:10.1037/a0037737

Baumle Amanda K and Fossett, Mark (2005) Statistical Discrimination in Employment: Its Practice, Conceptualization, and Implications for Public Policy. *The American Behavioral Scientist*. 49: 1250-1274.

Baumle, Amanda K. (2009) The demography of sexuality and the labor market. *International Handbook on the Demography of Sexuality*. 5, pp243-256.

Beauchamp, Tom L and Childress, James F (2008). *Principles of Biomedical Ethics*. (6th edition). New York: Oxford University Press.

Becker, Gary S (1957, 1971, 2nd ed.). *The Economics of Discrimination*. Chicago: University of Chicago Press.

Behaghel, Luc, Crepon, Bruno and Le Barbanchon, Thomas (2015) Unintended effects of anonymous resumes. *American Economic Journal: Applied Economics*. 7 (3): 1-27.

Benn, Stanley I (1978) Human rights – for whom and for what? In E Kamenka and AE-S Tay (Eds) *Human Rights*. London: E Arnold.

Bertrand, Marianne and Duflo, Esther (2016) Field experiments on discrimination. NBER Working Paper No. w22014, February 2016.

Bertrand, Marianne and Mullainathan, Sendhil (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94, 991-1013.

Bertrand, Marianne, Chugh, Dolly and Mullainathan, Sendhil (2005) Implicit Discrimination. *American Economic Review*, 95, 94-98.

Bicchieri, Christina (2006). Philosophy: what is to be done? *Topoi*, 25(1-2): 21–23. doi:10.1007/s11245-006-0025-y

Bichsel, Jacqueline (2012) Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations (Research Report). Louisville, CO: EDUCAUSE Center for Applied Research, August 2012. Retrieved from http://www.educause.edu/ecar.

Bloom, Dan (2006) Employment-focused programs for ex-prisoners: What have we learned, what are we learning and where should we go from here? MDRC Discussion paper, May 24, 2006. Retrieved from http://files.eric.ed.gov/fulltext/ED493008.pdf+A148

Blume, Lawrence E (2006) The Dynamics of Statistical Discrimination. *The Economic Journal*, 116: F480–F498.

Bourget, David and Chalmers, David J. (2013). What do philosophers believe? *Philosophical Studies*, 1–36. doi:10.1007/s11098-013-0259-7

Bryson, Jane, James, Susan and Keep, Ewart (2013) Recruitment and selection. In Stephen Bach and Martin Edwards (Ed.s) *Managing Human Resources: Human resource management in transition*, (Ch. 7, pp125-149), John Wiley & Sons Ltd.

Budig, Michelle J and England, Paula (2001) The Wage Penalty for Motherhood. *American Sociological Review,* 66(2): 204-225. Retrieved September 2016 from http://www.jstor.org/stable/2657415

Bunn William B. III, Stave, Gregg M, Downs, Kristen E, Alvir, Jose, Dirani, Riad (2006) Effects of smoking status on productivity loss. *Journal of Occupational and Environmental Medicine*, 48 (10): 1099-1108. DOI: 10.1097/01.jom.0000243406.08419.74 Retrieved from http://journals.lww.com/joem/Abstract/2006/10000/Effect_of_Smoking_Status_on_Productivity_Loss.18.aspx

Campbell, John, deBlois Peter, and Oblinger, Diana (2007) Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 41–57.

Canadian Superior Courts Judges Association (n.d.) The role of the courts. Retrieved from http://www.cscja-acjcs.ca/role_of_courts-en.asp?l=4

Cass, B. and Garde, P. (1984) Unemployment in the Western Region of Sydney: Job Seeking in a Local Labour Market. *SWRC Reports and Proceedings*. 47: 3–43.

Chadwick, Karen L (2006) Is leisure-time smoking a valid employment consideration? *Albany Law Review*, 70,117-141.

Chadwick, Ruth F, & Schroeder, Doris (2002). *Applied Ethics: Critical Concepts in Philosophy*. Taylor & Francis.

Colyvan, Mark, Regan, Helen, & Ferson, Scott (2001) Is it a crime to belong to a reference class? *Journal of Political Philosophy*, 9 (2), 168-181.

Correll, Shelley, Benard, Stephen and Paik, In (2007) Getting a job: is there a motherhood penalty? *American Journal of Sociology*, 112 (5): 1297-1339.

Courts of New Zealand (n.d.) The role of the courts. Retrieved from

   https://www.courtsofnz.govt.nz/about/system/role/overview

Cragg, Wesley (1992) *The Practice of Punishment: Toward a Theory of Restorative Justice*.

   Routledge, London and New York.

Crawford, Kate, & Schultz, Jason (2013) Big data and due process: Toward a framework to redress

   predictive privacy harms. NYU School of Law. Public Law Research Paper No. 13-64, 93-128.

Denny, Kathleen E. (n.d.) Workplace evaluations of parents by race: Unraveling perceptual

   penalties and premiums. Retrieved July 2016 from

   http://www.tessexperiments.org/data/Denny366.pdf

Dickinson, David L and Oaxaca Ronald L (2009) Statistical discrimination in labor markets: An

   experimental analysis. *Southern Economic Journal*,76 (1) : 16-31

Dittmer, Joel (2013). Ethics, Applied. *Internet Encyclopedia of Philosophy*. Retrieved from

   http://www.iep.utm.edu/ap-ethic/

Doleac, Jennifer and Hansen, Benjamin (2016) Does "Ban the box" help or hurt low-skilled

   workers? Statistical discrimination and employment outcomes when criminal histories are

   hidden. NBER Working Paper 22469, July 2016.

Duell, Dominik and Landa, Dimitri (30 March, 2015) Determinants of discrimination in a strategic

   setting. Working paper. Retrieved September 2016 from

   http://dominikduell.com/papers/ATT_paper.pdf

Edmonds, David (2006) *Caste Wars: A philosophy of discrimination*. London & New York:

   Routledge.

Eidelson, Benjamin (2013) Treating people as individuals. In Deborah Hellman & Sophia Moreau

   (Eds.) *The Philosophical Foundations of Discrimination Law* (pp.354-395). Oxford: Oxford

   University Press.

Enoch, David & Fisher, Talia (2015) Sense and "Sensitivity": Epistemic and instrumental approaches

   to statistical evidence. *Stanford Law Review*, 67 (3): 557-611.

Ewens M, Tomlin B and Wang LC (2011) Statistical discrimination or prejudice? A large sample field

   experiment. http://ewens.tepper.cmu.edu/wp-content/uploads/2011/04/statistical-

   discrimination2011.pdf (accessed 30 June 2011)

Fadlon, Yariv D (2015) Statistical discrimination and the implications of employer-employee race

   matches. *Journal of Labour Research*, 36: 232-248. DOI 10.1007/s12122-015-9203-2

Fang, Hanming and Moro, Andrea (2011) Theories of statistical discrimination and affirmative action: a survey. In Benhabib, J, Jackson MO and Bisin A. (Eds), Handbook of Social Economics, 1A, 133-200.

Federal Trade Commission (July 24, 2007). FTC releases report on effects of credit-based insurance scores. Retrieved 26 July 2016 from https://www.ftc.gov/news-events/press-releases/2007/07/ftc-releases-report-effects-credit-based-insurance-scores

Feltovich, Nick and Papageorgiou, C (2004) An experimental study of statistical discrimination by employers. *Southern Economic Journal*, 70, 837-849.

Fernandez, Roberto M and Greenberg, Jason (2013), Race, Network Hiring, and Statistical Discrimination. In Steve Mcdonald (Ed.) *Networks, Work and Inequality* (Research in the Sociology of Work, Volume 24), Emerald Group Publishing Limited, Pp.81-102.

Friedman, Milton (1953) The methodology of positive economics. In *Essays In Positive Economics*. Chicago: University of Chicago Press.

Gandy, Oscar H (2010) Engaging rational discrimination: exploring reasons for placing regulatory constraints on decision support systems. *Ethics and Information Technology*, 12 (1), 29-42.

Glauber, Rebecca (2007) Marriage and the motherhood wage penalty among African Americans, Hispanics, and Whites. *Journal of Marriage and Family*, 69: 951–961. doi:10.1111/j.1741-3737.2007.00423.x

Guerra-Pujol, Enrique (2014). Visualizing probabilistic proof, 7 *Washington University Jurisprudence Review* 39. Available at http://openscholarship.wustl.edu/law_jurisprudence/vol7/iss1/6

Hallevy, Gabriel (2013) *The right to be punished: Modern doctrinal sentencing*, DOI 10.1007/978-3-642-32388-1_2, Berlin Heidelberg: SpringerVerlag

Hamilton, David L., Sherman, Steven J., Crump, Sara A., & Spencer-Rodgers, Julie (2009) The role of entitativity in stereotyping: Processes and parameters. In T. D. Nelson (Ed.) *Handbook of prejudice, stereotyping, and discrimination* (pp. 179-198). Philadelphia, PA: Psychology Press.

Hamlin, Alan, & Stemplowska, Zofia (2012). Theory, Ideal Theory and the Theory of Ideals. *Political Studies Review*, 10(1), 48–62. doi:10.1111/j.1478-9302.2011.00244.x

Harford, Tim (2008) The dangers of rational racism. In *The Logic of Life*, London: Abacus

Hatzenbuehler, Mark L, Phelan, Jo C and Link, Bruce G. (2013) Stigma as a fundamental cause of population health inequalities. *American Journal of Public Health*, 103 (5): 813-821. Retrieved August 2016 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3682466/

Heath, Joseph (2014) Reasonable restrictions on underwriting. In I(pp.345-372) Oxford: Oxford University Press.

Heimer, Carol A (2002) Insuring more, ensuring less: The costs and benefits of private regulation through insurance. In Tom Baker and Jonathan Simon (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (pp116-145) Chicago and London: University of Chicago Press.

Hellman, Deborah (2008) *When is discrimination wrong?* Cambridge, Mass., USA: Harvard University Press.

Henry, Jessica & Jacobs, James (2007) Ban the box to promote ex-offender employment. Criminology and Public Policy, 6 (4): 755-762. Retrieved 1 November 2016 from http://www.nelp.org/content/uploads/2015/03/Henry-Jacobs.BantheBox.article.Oct-07.pdf

Horta, Oscar (2010) Discrimination in terms of moral exclusion. *Theoria*, 76: 314-322.

Huffcut, AI and Youngcourt, SS (2007) Employment Interviews. In DL Whetzel and GR Wheatson (Eds), *Applied measurement: Industrial psychology in human resource management*. Mahwah, NJ: Lawrence Erlbaum.

Information Commissioner's Office (n.d.) Data protection principles. Retrieved from https://ico.org.uk/for-organisations/guide-to-data-protection/data-protection-principles/

International Working Group on Data Protection in Telecommunications (IWGDPT) (2014) Working paper on big data and privacy: Privacy principles under pressure in the age of big data analytics. Retrived June 26, 2015 from http://dzlp.mk/sites/default/files/u972/WP_Big_Data_final_clean_675.48.12%20%281%29.pdf.

Jia, Pengfei (2014) Using predictive risk modeling to identify students at high risk of paper non-completion and program non-retention at university. MBus thesis. Auckland University of Technology.

Johnson, Jeffrey A (2014) The Ethics of Big Data in Higher Education. *International Review of Information Ethics*. Retrieved from http://www.i-r-i-e.net/inhalt/021/IRIE-021-Johnson.pdf

Jussim, Lee, Cain, Thomas R, Crawford, Jarret T, Harber, Kent, Cohen, Florette (2009) The unbearable accuracy of stereotypes. In T. D. Nelson (Ed.) *Handbook of prejudice, stereotyping, and discrimination* (pp. 179-198). Philadelphia, PA: Psychology Press.

Kaas, Leo and Manger, Christian (2010) Ethnic discrimination in Germany's labour market: A field experiment. IZA Discussion Paper No. 4717, February 2010. Retrieved from http://ftp.iza.org/dp4741.pdf

Kaiser, Matthias, Millar, Kate, Thorstensen, Erik, & Tomkins, Sandy (2007). Developing the ethical matrix as a decision support framework: GM fish as a case study. *Journal of Agricultural and Environmental Ethics*, 20(1), 65–80. doi:10.1007/s10806-006-9023-8

Katz, Irwin (2014) *Stigma: A social psychological analysis.* New York: Psychology Press.

Kay, David, Korn, Naomi and Oppenheim, Charles (2012) Legal, Risk and Ethical Aspects of Analytics in Higher Education. *CETIS Analytics Series*, 1 (6): 1-30.

Klumpp, Tilman and Su, Xuejuan (2013) Second-order Statistical Discrimination. *Journal of Public Economics*, 97 (8): 108-116.

Koopman, Colin (2012) Pragmatist Resources for Experimental Philosophy: Inquiry in Place of Intuition. *Journal of Speculative Philosophy*, 26 (1): 1-24.

Kopelman, Loretta M. (1990) What is applied about 'Applied' philosophy? *Journal of Medicine and Philosophy*, 15(2): 199-218.

Kovacic, Zlatko (2010) Early Prediction of Student Success: Mining Students Enrolment Data. Proceedings of Informing Science & IT Education Conference (InSITE) 2010.

Lahey, Joanna N (2008) Age, women, and hiring. *The Journal of Human Resources*, 43 (1); 30-56.

Le H, Oh I-S, Shaffer, J and Schmidt, F (2007) Implications of methodological advances for the practice of personnel selection: How practitioners benefit from meta-analysis. *Academy of Management Perspectives*, 21 (3), 6-15.

LeBlanc, Greg (1995) Discrimination in the labor market. *Canadian Journal of Economics*, 28: 702-17.

Lecker, Martin J (2009) The smoking penalty: Distributive justice or smokism? *Journal of Business ethics*, 84: 47-64.

Levashina, Julia and Campion, Michael A (2009) Expected practices in background checking: Review of the Human Resource Management Literature. *Employee Responsibilities and Rights Journal*, 21: 231-49.

Levy, Jacob T. (2014) There's No Such Thing as Ideal Theory (SSRN Scholarly Paper No. ID 2420125). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2420125

Lippert-Rasmussen, Kasper (2006) Private discrimination: A prioritarian, desert-accommodating account. *San Diego Law Review*, 43: 817-856.

Lippert-Rasmussen, Kasper (2007) Nothing personal: On statistical discrimination. *Journal of Political Philosophy*, 15 (4): 385–403. DOI: 10.1111/j.1467-9760.2007.00285.x

Lippert-Rasmussen, Kasper (2011) "We are all different": Statistical discrimination and the right to be treated as an individual. *The Journal of Ethics*, 15: 47-59.

Lippert-Rasmussen, Kasper (2014) *Born free and equal? A philosophical inquiry into the nature of discrimination*. New York: Oxford University Press.

Lokken, Fred, & Mullins, C (2015) ITC 2014 Distance Education Survey Results. Retrieved November 2015 from http://www.itcnetwork.org/membership/itc-distance-education-survey-results.html

Long, Phillip, & Siemens, George (2011) Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), 31–40.

Lord Low of Dalston (16 May 2007, 4.45pm) Offender Management Bill. Lords Hansard Home Page. Parliamentary business, Publications and Records.   Retrieved from http://www.publications.parliament.uk/pa/ld200607/ldhansrd/text/70516-0004.htm

Lundberg, Shelly and Startz, Richard (1983) Private discrimination and social intervention in competitive labor markets. *American Economic Review*, 73: 340-7.

Maitzen, Stephen (1991) The Ethics of Statistical Discrimination. *Social Theory and Practice*, 17 (1): 23-45.

Masclet, David, Peterle, Emmanuel, and Larribeau, Sophie (2012) The role of information in deterring discrimination: A new experimental evidence of statistical discrimination. Center for Research in Economics and Management, Working Paper 2012-28. Retrieved from http://crem.univ-rennes1.fr/wp/2012/201238.pdf

Mayer-Schönberger, Viktor and Cukier, Kenneth (2014) *Learning with big data*. Boston/New York: Houghton Mifflin Harcourt.

McBain, James (2012) Issue Introduction. *Essays in Philosophy*, 13, 1, Article 1.

McKinlay, Steve (2015) The ethics of data convergence and predictive modelling. Paper presented at Australian Association of Professional and Applied Ethics Conference, 9-12 July, University of Auckland, New Zealand.

McSweeney, Tim and Hough, Mike (2006) Supporting offenders with multiple needs: Lessons for the 'mixed economy' model of service provision. *Criminology and Criminal Justice*, 6: 107-125. DOI: 10.1177/1748895806060669

Meyers, Glenn (1985) An analysis of experience rating. *PCAS*. Retrieved 28 October 2016 from http://www.casact.org/pubs/proceed/proceed85/85278.pdf

Mialon, Sue H and Yoo, Seung Han (2014) Employers' preference for discrimination. Emory Economics Working Paper Series 1402. Retrieved from http://economics.emory.edu/home/assets/workingpapers/smialon_14_02_paper.pdf

Moreau, Sophia (2010) Discrimination as Negligence. *Canadian Journal of Philosophy*, 40:sup1, 123-149.

Moreau, Sophia (2013) In Defense of a Liberty-based Account of Discrimination. In Deborah Hellman & Sophia Moreau (Ed.s) *The Philosophical Foundations of Discrimination Law* (Pp.71-86). Oxford: Oxford University Press.

Moro, Andrea and Norman, P (2003) Empirical Implications of Statistical Discrimination on the Returns to Measures of Skills. *Annales d'Economie et de Statistique*, 71-72: 399-417.

Mussel, Patrick (2012) Introducing the construct curiosity for predicting job performance. *Journal of Organizational Behavior*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/job.1809/full

Nunn, Alex, Bickerstaff, Tim, Hogarth, Terence, Bosworth, Derek, Green, Anne E and Owen, David (2010) Postcode Selection? Employers' Use of Area- and Address-based Information Shortcuts in Recruitment Decision. Department for Work and Pensions Research Report 664.

OAAI, Marist College (2012) Open Academic Analytics Initiative. Retrieved November 2015 from https://confluence.sakaiproject.org/pages/viewpage.action?pageId=75671025

Oblinger, Diana (2012) Let's Talk … Analytics. *EDUCAUSE Review*, 47(4): 10–13.

Open University (Oct 2014) Ethical use of student data for learning analytics policy FAQs. Retrieved October 2015 from http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-student-data-faq.pdf

Open University (Sep 2014) Policy on ethical use of student data for learning analytics. Retrieved October 2015 from http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-use-of-student-data-policy.pdf

Pachankis, John E (2007) The psychological implications of concealing a stigma: A cognitive-affective behavioral model. *Psychological Bulletin*, 133 (2): 328-345.

Palmer, Daniel E (2007) Insurance, risk assessment and fairness: An ethical analysis. *Insurance ethics for a more ethical world, Research in Ethical Issues in Organisations*, 7: 113-126.

Pardo, Abelardo and Siemens, George (2014) Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45: 438-450.

Pedulla, David (2014) The Positive Consequences of Negative Stereotypes: Race, Sexual Orientation, and the Job Application Process. *Social Psychology Quarterly*, 77 (1). doi: 10.1177/0190272513506229

Phelps, Edmund S (1972) The statistical theory of racism and sexism. *American Economic Review*, 62: 659-61.

Pinquet, Jean (2012) Experience rating in non-life insurance. Ecole Polytechnique, Centre National de la Recherche Scientifique, Cahier no 2012-10. Retrieved 28 October 2016 from https://hal.archives-ouvertes.fr/file/index/docid/677100/filename/cahier_de_recherche_2010-12.pdf

Powell, David and Goldman, Dana (2016) Disentangling Moral Hazard and Adverse Selection in Private Health Insurance. NBER Working Paper N0. 21858. Accessed 29 June 2016: http://www.nber.org/papers/w21858

Prinsloo, Paul and Slade, Sharon (2014) Educational triage in open distance learning: Walking a moral tightrope. *The International Review of Research in Open and Distance Learning*. Retrieved from http://oro.open.ac.uk/40903/1/3007

Rawls, John (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Richards, Neil and King, Jonathan (2014) Big data ethics. *Wake Forest Law Review*, 49: 393–432.

Richardson, Michelle, Abraham, Charles, and Bond, Rod (2012) Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin*, 138(2): 353–387. DOI: 10.1037/a0026838

Rissing, Den A and Castilla, Emilio J (2014) House of Green Cards: Statistical or preference-based inequality in the employment of foreign nationals. *American Sociological Review*, 79 (6): 1226-1255. Retrieved from http://asr.sagepub.com/content/79/6/1226

Robroek, Suzan JW, van den Berg, Tilja IJ, Plat, Jan F, Budorf, Alex (2011) The role of obesity and lifestyle behaviours in a productive workforce. *Occupational & Environmental Medicine*, 68: 134-139. DOI: 10.1136/oem.2010.055962

Roehling, MV (2002) Weight discrimination in the American workplace: Ethical issues and analysis. *Journal of Business Ethics*, 40: 177-89.

Rossler, Beate (2005) *The Value of Privacy*. Cambridge, UK: Polity Press

Sackett, Paul R and Lievens, Filip (2008) Personnel Selection. *Annual Review of Psychology*, 59: 419-50.

SAS Institute Inc, (n.d.) Who's using it? Machine learning: What it is and why it matters. [webpage] http://www.sas.com/en_us/insights/analytics/machine-learning.html#machine-learning-users

Sattinger, Michael (1998) Statistical Discrimination with Employment Criteria. *International Economic Review*, 39: 205-30.

Schauer, Frederick (2003) *Profiles, probabilities and stereotypes*. Cambridge, Massachusetts: Harvard University Press.

Schick, Andreas and Steckel, Richard H (2010) Height as a Proxy for Cognitive and Non-cognitive Ability. NBER Working Paper No. 16570.

Schmidt, Frank L and Hunter, John E (1998) The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin*, 124: 262-74.

Schneebaum, Alyssa (2013) Motherhood and the lesbian wage premium. University of Amherst-Massachusetts Working paper No. 4. Retrieved from https://www.umass.edu/economics/publications/2013-04.pdf

Schoeman, Ferdinand (1987) Statistical vs direct evidence. *Nous*, 21 (2): 179-198.

Schoeman, Ferdinand D (Ed). (1984) Privacy: Philosophical dimensions of the literature. In *Philosophical Dimensions of Privacy: An Anthology*. Cambridge: Cambridge University Press

Scholes, Vanessa (2014) You are not worth the risk: Lawful discrimination in hiring. *Rationality, Markets and Morals*, 5: 13-29.

Scholes, Vanessa (2016) Risky business – the ethics of judging individuals based on group statistics. In Marco Grix and Tim Dare (Eds.) *Contemporary Issues in Applied and Professional Ethics* (pp. 169–188). Emerald. doi:10.1108/s1529-209620160000015010

Scholes, Vanessa (2016) The ethics of using learning analytics to categorize students on risk. *Educational Technology Research & Development*. DOI: 10.1007/s11423-016-9458-1

Schwab, S (1986) Is Statistical Discrimination Efficient? *The American Economic Review*, 76: 228-34.

Sclater, Niall (2014a) Code of practice for learning analytics: A literature review of the ethical and legal issues. JISC, London.

Sclater, Niall (2014b) Learning analytics: The current state of play in UK higher and further education. JISC. Retrieved from http://repository.jisc.ac.uk/5657/1/Learning_analytics_report.pdf

Sclater, Niall and Bailey, Paul (2015) Code of practice for learning analytics. JISC, London. Retrieved October 2015 from https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics

Segall, Shlomi (2012) What's so bad about discrimination? *Utilitas*, 24(1): 82-100. doi:10.1017/S0953820811000379

Sherman, Bruce W and Lynch, Wendy, D (2013) The relationship between smoking and health care, workers' compensation and productivity costs for a large employer. *Journal of Occupational and Environmental Medicine*, 55 (8): 879-884. DOI: 10.1097/JOM.0b013e31829f3129

Simpson, Ormond (2009) Open to people, open with people: Ethical issues in open learning. In Ugur Demiray and Ramesh C. Sharma (Ed.s) *Ethical Practices and Implications in Distance Learning* (pp.199-215). New York, USA: Information Science Reference.

Slade, Sharon, & Prinsloo, Paul (2013) Learning Analytics: ethical issues and dilemmas. *American Behavioural Scientist*, 57 (10): 1509-1528.

Stafford, Mark C and Scott, Richard, R. (1986) Stigma, Deviance and Social Control. The Dilemma of Difference, *Perspectives in Social Psychology*, pp77-91.

Stangor, Charles (2009) The study of stereotyping, prejudice, and discrimination in social psychology: A quick history of theory and research. In T. D. Nelson (Ed.) *Handbook of prejudice, stereotyping, and discrimination* (pp3-22). Philadelphia, PA: Psychology Press.

Stemplowska, Zofia and Swift, Adam (2013) Rawls on ideal and nonideal theory. In Jon Mandle and David Reidy (Ed.s) *A Companion to Rawls*. (pp112-127) West Sussex, UK: Wiley Blackwell.

Stone, Deborah (2002) Beyond moral hazard: Insurance as moral opportunity. In Tom Baker and Jonathan Simons (Ed.s) *Embracing Risk: The changing culture of insurance and responsibility*. (Pp52-79). Chicago and London: University of Chicago Press.

Stuber, Jennifer, Meyer, Ilan and Link, Bruce (2008) Stigma, prejudice, discrimination and health. *Social Science & Medicine*, 67 (3): 351-357.

Taylor, Paul, Keelty, Yvonne and McDonnell, Bridget (2002) Evolving Personnel Selection Practices in New Zealand Organisations and Recruitment Firms. *New Zealand Journal of Psychology*, 31: 8-18.

Thompson, Paul and Smith, Chris (2009) Waving, not drowning: Explaining and exploring the resilience of labor process theory. *Employee Responsibilities and Rights Journal*, 21: 253-62.

Thomsen, Frej Klem (2013) But some groups are more equal than others: A critical review of the group criterion in the concept of discrimination. *Social Theory and Practice*, 39 (1): 120-146.

Truth and Justice Commission (2011) Volume One Report of the Truth and Justice Commission. URL: http://www-previous.usip.org/files/ROL/TJC_Vol1.pdf [29.07.2013].

Tyler, Tom R (2014) How do the courts create popular legitimacy?: The role of establishing the truth, punishing justly, and/or acting through just procedures. Faculty Scholarship Series. Paper 4991. Retrieved from http://digitalcommons.law.yale.edu/fss_papers/4991

Vallentyne, Peter (2006) Left libertarianism and private discrimination. *San Diego Law Review*, 43: 981-994

Vuong, Annalies, Nixon, Tristan, & Towle, Brendan (2011) A method for finding prerequisites in a curriculum. From M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero & J Stamper (Eds.) *Proceedings of the 4th International Conference on Educational Data Mining*. Retrieved from http://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm2011_paper8_short_Vuong.pdf

Wagner, Ellen, & Hartman, Joel (2013) Welcome to the Era of Big Data and Predictive Analytics in Higher Education. SHEEO (State Higher Education Executive Officers Association Higher Education Policy Conference 2013 [presentation]
Retrieved from: http://www.sheeo.org/sites/default/files/0808-1430-plen.pdf

Wang, Xuan (2011) Characteristics of the human resource executives and companies that use more sophisticated employee selection methods. Thesis, Emporia State University. Retrieved 25 July 2016 from https://esirc.emporia.edu/bitstream/handle/123456789/998/XuanWang%20thesis.docx?sequence=1

Wertheimer, Alan (2006) Reflections on discrimination. *San Diego Law Review*, 43: 945-980

Willis, James & Pistilli, Matthew (April 7, 2014) Ethical discourse: Guiding the future of learning analytics. *EDUCAUSE Review*. Retrieved October 2015 from http://www.educause.edu/ero/article/ethical-discourse-guiding-future-learning-analytics

Willis, James (August 25, 2014) Learning analytics and ethics: A framework beyond utilitarianism. *EDUCAUSE Review*. Retrieved October 2015 from http://er.educause.edu/articles/2014/8/learning-analytics-and-ethics-a-framework-beyond-utilitarianism

Witt, Philip, Dattilio, Frank and Bradford, John (2011) Sex Offender evaluations. In E. Drogin, F. M. Dattilio, R.L. Sadoff & T. G. Gutheil (Ed.s) *Handbook of Forensic Assessment: Psychological and Psychiatric Perspectives*. Wiley Online.

Woodley, Alan and Simpson, Ormond (2014) Student Dropout: The elephant in the room. In Olaf Zawacki-Richter & Terry Anderson (Ed.s) *Online Distance Education: Towards a research agenda*. (Pp.459-483). Edmonton, Canada: AU Press.

Yip, Chi Man and Wong, Raymond Sin-Kwok (2014) Gender-oriented statistical discrimination theory: Empirical evidence from the Hong Kong labor market. *Research in Social Stratification and Mobility*, 37: 43-59.

Zwolinski, Matt (2006) Why not regulate private discrimination? *San Diego Law Review*, 43: 1043-1061.