Journal of Practical Ethics

CONTENTS

The Journal of Practical Ethics is available online, free of charge, at:
http://jpe.ox.ac.uk

Editorial Policy
The *Journal of Practical Ethics* is an invitation only journal. Papers are anonymously appraised prior to publication by expert reviewers who are not part of the editorial staff. It is entirely open access online, and print copies may be ordered at cost price via a print-on-demand service. Authors and reviewers are offered an honorarium for accepted articles. The journal aims to bring the best in academic moral and political philosophy, applied to practical matters, to a broader student or interested public audience. It seeks to promote informed, rational debate, and is not tied to any one particular viewpoint. the journal will present a range of views and conclusions within the analytic philosophy tradition. It is funded through the generous support of the *Uehiro Foundation in Ethics and Education.*

The Feminist Argument Against Supporting Care

ANCA GHEAUS

ABSTRACT

Care-supporting policies incentivise women's withdrawal from the labour market, thereby reinforcing statistical discrimination and further undermining equality of opportunities between women and men for positions of advantage. This, I argue, is not sufficient reason against such policies. Supporting care also improves the overall condition of disadvantaged women who are care-givers; justice gives priority to the latter. Moreover, some of the most advantageous existing jobs entail excessive benefits; we should discount the value of allocating such jobs meritocratically. Further, women who have a real chance to occupy positions of advantage have most likely already enjoyed more than their fair share of opportunities; they lack a claim to more. Women can have a complaint grounded in the expressive disvalue of sexist discrimination. This gives them special claims against men occupying the vast majority of top positions and against their higher share of opportunities for positions of advantage. But their claim does not speak against care-supporting policies.

INTRODUCTION

Feminist scholars and activists regularly defend public policies that support care-giving and care-givers. Some of these policies support care directly, by entitling to financial compensation individuals who provide care to their family members and who thereby often forego employment or downsize their career ambitions. Examples include a mandate that employers split the paycheck between their employee and their stay at home partner (Okin, 1989), or a 'caregiver's allowance' in the form of a

voucher given to parents who provide hands-on care (Alstott, 2004). Other measures are meant to support the combination of care-giving and paid work, and include paid and relatively long parental leaves, flexible working hours for parents, better compensated part-time work and subsidies for non-parental childcare (Williams, 2000; Gornik and Meyers, 2003; Schouten, 2019). Finally, some policies whose aims are more general—such as a universal basic income—would also support-caregivers (Elgarte, 2008; Baker, 2008). These diverse policies are justified by different considerations. Yet most[1] of them trigger a general feminist worry: that protecting care-givers unavoidably provides women with incentives to opt out of the labour market at higher rates than they already do, and at much higher rates than men. This, the criticism goes, will reinforce statistical discrimination and sexist stereotypes that portray women as less well-performing employees, and thus further undermine equality of opportunities between women and men for positions of advantage.

This criticism, I argue, doesn't generate a powerful reason against care-supporting policies—not in general, and even less so in unjustly unequal societies. One broad argument is as follows: even if it is true that supporting care makes it more difficult for women to access positions of advantage compared to men, it is also likely to *improve* the overall condition of women who are most disadvantaged by their care-giving social role. Not all the above policy proposals would benefit the worse off women equally, but all would do so considerably. At the same time, it is mostly better off women who have an interest in equal opportunities to positions of advantage. It is highly plausible that the interests of the worst off ought to be given priority. Here I will not engage in any detail with this general argument but focus, instead, on the particularly strong case in favour of supporting care-giving in unjust societies.

I assume that existing societies fail to observe any plausible principle of distributing outcomes. I don't commit here to a particular theory of justice; such a principle may be luck-egalitarian, permitting only inequalities that track individual responsibility, and therefore deeming as unjust cases where some people are worse off than others out of no fault or choice of their own. Or it can be a prioritarian principle, requiring that the worse off an individual is, the more weight we give to her interests. Or it can be a Rawlsian principle, allowing only for inequalities that are necessary to

---

1. But not the subsidising of non-parental care, which of all the listed policies, would actually enable women to compete with men as full-time employees.

improve the prospects of the worst-off members of society.[2] Finally, much of my analysis should also appeal to sufficientarians, who rule out inequalities to the extent to which they leave some individuals under a threshold of sufficiency. Societies which fail to conform to any of these principles, or alternative principles of justice, I will call "unjust societies". In unjust societies, the better off women lack a claim of justice with respect to many of the most advantageous jobs. To this conclusion I advance two arguments. The first applies to the very best jobs, which I call, by stipulation, "top positions". This is a placeholder term, but plausible candidates include being the CEO of a large company, the best paid jobs in the entertainment industry etc. The second argument applies to positions of advantage in general, plausible examples of which are jobs in law, academia, medicine etc. Given its focus on conflicts of interests between worse off and better off women, this paper echoes the current popular debates between a feminism interested in the fate of poor and marginalised women and so-called "boardroom feminism". While I cannot engage directly with this debate here, my overall argument provides support to the former. I rely on a common distinction between equality of outcome—concerning where people end up with respect to, for instance, income, social status, or power—and equality of opportunity for such outcomes—that is, a level-playing field in competing for the outcomes.

Argument 1

- P1 Nobody has a right to a position that confers unjust advantages on those who hold it.

- P2 We ought to discount the value of equality of opportunity for positions to which nobody has a right.

- P3 Many of the top positions in societies with large unjustifiable inequalities of outcome confer unjust advantages on those who hold them.

- C1 Therefore, we ought to discount the value of equality of opportunity for many of the top positions in societies with large unjustifiable inequalities of outcome.

There is a legitimate question about the relationship between the magnitude of

---

2.   In Rawls' overall theory of justice this principle determines which social positions should exist. Rawls believes that access to these positions should be governed by a principle of fair equality of opportunities (discussed below.)

unjustifiable inequalities attached to a position and the plausibility of P2 with respect to that position. One possibility is that the more modest the deviation from justice, the less we should discount the value of distributing the position in question according to equality of opportunity. So, for instance, if the pay-offs attached to a position deviate only marginally from what justice permits, it is almost as desirable to distribute it according to equality of opportunity as it is in the case of positions that carry permissible rewards. I cannot address this problem properly here, but I assume that at least some top positions in existing societies involve benefits that are entirely out of proportion with the demands of distributive justice, and that we should attach almost no value to distributing those positions meritocratically, that is according to equality of opportunity.

Argument 2

- P4 Most people who have a shot at positions of advantage are likely to already have had more than their fair share of opportunities for such positions.

- P5 People lack a claim to more than their fair share of opportunities for positions of advantage.

- C2 Therefore, most people who have a shot at positions of advantage do not have a claim to more opportunities than they already have for such positions.

According to several feminists, the fact that a policy worsens the inequalities of opportunities between women and men to positions of advantage is a serious consideration against them (Robeyns, 2000; Robeyns, 2001; Gheaus, 2008), perhaps even a defeating one (Baber, 1990; Bergmann, 1998; Bergmann, 2008). Against these views, Arguments 1 and 2 show that all women lack a claim to top positions and most women who stand a chance for positions of advantage lack a claim to more opportunities for them than they already have. But all women *do* have special claims against men having more opportunities for them, including for top positions: this is a complaint grounded in the *expressive* disvalue of denying women the same opportunities that men have—as I explain below. The complaint provides grounds for implementing policies that combat gender stereotypes which associate women with care and present them as less desirable workers than men. In particular, the expressive complaint supports policies implementing gender parity in the distribution of posi-

tions of advantage. Such parity is desirable both for its immediate expressive value and, potentially, for its long-term consequences in breaking the association between women and domesticity on the one hand, and between men and paid employment on the other hand (Radcliffe-Richards, 2014). Yet one need not aim at the much more demanding ideal of equality of opportunities in order to address women's expressive complaint.

I assume that, given the large unjustifiable inequalities of both outcome and of opportunities in contemporary societies, most positions where discrimination against women would increase due to care-supporting policies are likely to be like the positions in arguments 1 and 2. This is not a purely speculative point: the main feminist complaint against care-supporting policies is that they will increase a kind of discrimination against women which, as I explain below, is a lot more likely to affect access to positions of advantage. This is not to deny that other kinds of sexist discrimination apply also, or especially, to women in middle and lower grade jobs. But this is not the kind of discrimination likely to be enhanced by care-supporting policies. Therefore, care-supporting policies are significantly less likely to increase inequality of opportunities between women and men for middle-grade jobs and even less so for lower-grade jobs. Moreover, with respect to the lowest-grade positions, such as menial work, the disvalue of a great many of these positions may also mean that opportunities for them are less important to individuals likely to occupy them than receiving care-giving support. Therefore, assuming that a care-supporting policy is otherwise justified, egalitarian feminists have much better reason to welcome than to resist it, even if it sets back women's opportunities to positions of advantage.

Equal opportunities criticism against supporting care has been raised in connection with all the above-mentioned policies (Baber, 1990; Bergmann, 2008; Miller, 2016). But the most articulated debate to date concerns the introduction, in existing developed societies, of a universal, unconditional basic income (henceforth "BI")—that is, a regular income financed through taxation and paid to all the members of a political community regardless of any other sources of income and independent from the recipients' productive participation in social cooperation and of their level of financial need. Here I refer to the feminist debate concerning BI merely to illustrate the tradition of appealing to equality of opportunity for positions of advantage as a weighty reason against its introduction. All care-supporting policies are open to this criticism, since they are likely to activate the same mechanisms that incentivise the perpetuation of the gendered division of labour. I do not aim, as such, to defend

any of these policies, the overall evaluation of which—especially in the case of BI—depends on many normative and empirical considerations, including matters of feasibility and existing alternatives. My aim is to assess, and reject, a particular argument against them.

The next section offers some theoretical clarifications; the third presents the feminist debate about the care-supporting effects of BI. In the fourth section I elaborate on Arguments 1 and 2 and their implications. I end with brief conclusions.

BACKGROUND CLARIFICATIONS: FEMINIST EGALITARIANISM

There is broad consensus that equality must govern the access to some goods, such as political and civil rights, and that avoidable inequalities in the distribution of other goods require justification. Examples of the latter include resources (like money), welfare (that is, how well one's life goes), and capabilities (for instance, being able to move freely), and, possibly, relational goods such as social inclusion and non-stigmatisation.[3] Some philosophers justify departures from equality in the distribution of such goods in cases where these departures result from individuals' autonomous choices; others believe departures from equality are permissible if they are necessary in order to make the worst off as well off as possible. It is not obvious that any of these versions of egalitarianism—widely understood to include prioritarianism—involves, as such, any particular interest in whether more women than men are the victims of injustice, or in whether gender norms in particular (as opposed to other unjustified mechanisms) are responsible for inequalities. For instance, being a welfarist luck egalitarian—that is, endorsing equality of welfare—does not as such give you reason to find it particularly objectionable if women are disproportionately affected by unchosen inequalities in welfare due, for instance, to the gendered division of labour; on this view, what matters are inequalities in welfare between individuals. Being a prioritarian does not as such give you reason to object to the feminisation of poverty. And being a relational egalitarian—believing that justice requires that individuals relate to each other as equals—does not

---

3.   There is a debate on whether relational goods such as inclusion, non-stigmatisation or non-domination can be part of the metric of distributive justice. I assume they can, and argued so in "Hikers in Flip-Flops. Luck Egalitarianism, Democratic Equality and the Distribuenda of Justice" (2018). Taking a stance in this debate bears on the form, and scope, but not on the substance, of the present argument.

as such give you reason to be particularly concerned with women's social marginalisation or stigmatisation. "As such", because many egalitarians do in fact object to gendered inequalities; but, I assume, they do so in virtue of additional beliefs—for instance, concerning the role that a history of discrimination and oppression plays in explaining such inequalities.

By contrast, to be a feminist *is* to be interested in whether the group of those affected by injustice is gendered, at least to the extent to which inequalities can be attributable to gender norms. Feminism, too, comes in several varieties. A common denominator is the joint belief that (a) women and men are moral equals and therefore (b) women and men should have equal opportunities, at each stage of life, to access those goods the distribution of which is relevant to justice. Many feminists also endorse the view that (c) women should not fare worse than men, on the relevant conception of justice, due, even in part, to their sex. The last point goes beyond the request for equal opportunities: it states that women should not be worse off merely because they have different native competencies than men (if they do). So, for instance, if justice allows different compensation for different skills, and women end up being overly represented amongst the worse off because they tend to have native skills that receive worse compensation on the market, feminists may object to the fact that pay-offs are attached to marketable skills in a way that disadvantages women. Depending on their particular conception of justice, different feminists will identify different inequalities as objectionable. For instance, to qualify as feminist, a libertarian must object to women suffering worse rights violations than men; sufficientarian feminists must object to the feminisation of poverty, etc.

Egalitarian feminists' interest in distributive justice is at least in part fueled by a sense that in many societies women constitute the majority of those denied their fair share—mainly in terms of wealth, status and power—and that this is due, in part, to the fact that they are women. This complaint goes beyond a complaint of unequal opportunities to occupy positions of advantage. If women and men had the same opportunities—in the robust sense specified at (b) above—but more women than men ended up as poor, marginalised etc., egalitarian feminists could have a double complaint. The first is that, in virtue of being egalitarians, they deem poverty unjust. But if this state of affairs is due to background economic rules and expectations that relegate care-givers to poverty, they also have grounds to complain in virtue of being feminists, even if women end up in caring positions *via* a distribution of work which is governed by equality of opportunity.

As specified by point (b) above, feminists believe that women and men should be able to compete unencumbered by socially created obstacles related to them being women or men. Some, such as Janet Radcliffe-Richards (2014), believe that this requirement fully captures feminist demands concerning the distribution of positions of advantage. I don't deny that equality of opportunity between women and men at each stage of life is a worthy, and weighty, goal. But— as the arguments of this paper show—in societies that harbour large unjustified inequalities of outcome we should discount this widely-held ideal of gender justice when applied to positions of advantage. In such circumstances, feminists should, instead, strive to eliminate the feminisation of the worse off groups.

THE FEMINIST ARGUMENT AGAINST CARE-SUPPORTING POLICIES

Feminists have always been divided on the merits of care-supporting policies with respect to how they would affect women and gender equality. As already indicated, at least some of the disagreement is principled: such policies are likely to significantly improve, on the whole, the material situation, power and status of the worse off women, and in particular care-givers, while at the same time making it harder to achieve the goal of equal opportunities between women and men for the best jobs.

The debate about BI provides a detailed illustration of this feminist conundrum. BI would help lift from poverty women in jobs with low pay, few benefits and little security. If generous enough to provide financial independence, BI would provide exit to women who are in abusive domestic relationships and economically dependent on their partners. It would also protect from old age poverty women who have not joined the labour market (or left it prematurely) and who therefore do not qualify for welfare benefits (Whithorn, 2013; Robeyns, 2000; Elgarte, 2008). It would therefore benefit women who find themselves in these circumstances due to their caring responsibilities. At the same time, BI is likely to have a negative effect on women's opportunities in the labour market and, possibly, on their access to positions of power and prestige (Robeyns, 2000; Gheaus, 2008).[4] One of the earliest feminist concerns was that BI will diminish women's participation in the labour market by making it affordable for more women to opt out of paid work—fully or partially—and instead

---

4.   I cannot evaluate here the accuracy of these predictions, nor is this necessary for the sake of a normative analysis of the argument from equal opportunities.

do care work in the family (Orloff, 2013).[5] Having withdrawn from paid work for a period and "specialised" in care work, it then becomes more difficult for women to access positions of advantage when they re-enter the labour market; more likely, such women will continue to be mainly care-providers or stuck in undesirable jobs. The *status quo* is already bad. In Europe, for instance, in spite of doing well, and sometimes out-competing men, in higher education, women make up the vast majority of part-time workers. They are a lot less likely than men to hold a job if they have children under 12, have an overall lower employment rate, earn less, and represent only a small fraction of executive boards' membership or holders of top positions such as CEO (Eurostat, 2017; European Commission, 2016). Women also tend to do more voluntary work and informal, unpaid, care work, which is usually low status. As a result, they earn less and have less decision-making power than men as well as lower social status.

Some philosophers think that feminists should welcome policies that *enlarge* women's opportunity sets even if as a result women will participate (even) less in the labour market. On this view, women's lower market attachment is a morally innocent effect because it reflects individual women's choices.[6] This argument, however, is too quick—as two feminist rejoinders show.

NOT ALWAYS THEIR (VOLUNTARY) *CHOICES*

The first feminist rejoinder is that, at least sometimes, women's decisions to prioritise caring for children, elderly parents or sick relatives do not carry the usual normative implications of autonomous choice because they are not made under the circumstances necessary for full autonomy. Women are (still) being socialised into assuming caring roles—with the help of gender-specific toys and bodily discipline, educational activities and generally by being pressured to comply with feminine social roles through active encouragements and/or through penalties. If so, perhaps women who choose part-time paid work or full-time home-making in order to look after family members in need for care don't enjoy the same degree of autonomy that

---

5.   It is worth noting that, by keeping women away from the job market, BI can in some cases set back their emancipation from abusive relationships. Employment gives people access to social networks. However, it does not do so uniquely: volunteering—for which BI has been said to provide incentives— has the same benefit.

6.   Moreover, in Philippe Van Parijs's words: "there is something insulting about considering women, in particular less-skilled women, to be less able than men to make a wise use of these expanded options." (Vanderborght and Widerquist, 2013, p. 142)

a person with a less gendered socialisation would enjoy in an otherwise identical choice context. It then seems unfair for such women to pay the full price of their choices in terms of loss of income, status and loss of other goods that one can only access through (good) jobs (Gheaus and Herzog, 2016). Policies are objectionable to the extent to which they sponsor imprudent choices for which the choice-makers lack full liability. The feminist worry that care-supporting policies will encourage some women to leave the job market in order to care more is indeed premised on the thought that this is can be an imprudent choice.

Recently, the gendered socialisation worry has taken one specific form, for which psychologists have coined a new term: stereotype threat. This is the fear that others will perceive you in light of negative stereotypes associated with the group with which you identify—in this case, gender—fear of which tends to inadvertently confirm the existing stereotypes (Steele and Aronson, 1995). There are reasons to think that women suffer from stereotype threat when performing within the traditional "male" territory and, as a result, perform worse than they would otherwise perform (and worse than men) precisely because they are expected to do so. This can explain, to some extent, women's lower achievements in fields dominated by, and traditionally associated with, men (Stricker and Ward, 2004).[7] It may also explain, in part, women's higher willingness to leave the job market. Again, a policy sponsoring an imprudent choice, choice whose causal history that includes (possibly many) episodes of non-autonomous choices is, to that extent, objectionable.

Finally, a woman's opting out of the labour market in order to care for someone in need of care can fail to qualify as a fully autonomous choice for an entirely different reason. Even someone who was free from any gendered socialisation, and unsusceptible to stereotype threat, could nevertheless have her autonomy curtailed because she finds herself unfairly saddled with a special duty of care. Individuals who become parents thereby acquire a duty of care towards their children; sometimes, foreseeably or not, one of the parents must switch to part-time work, or give up paid employment temporarily in order to care for the child, especially in societies where adequate daycare is unavailable or unaffordable. In cases when neither parent wants to do this, women, who are usually socialised as care-givers and who tend to earn less than men, face economic, and possibly psychological, pressure to be the ones who give

---

7. Stereotype threat research is not without critics; in particular, some raised doubts about the methodology of research that attributes women's and men's different achievement in mathematics to stereotype threat (Stoet and Geary, 2012).

up full employment (Okin, 1989, especially pp. 134-169; Allen, 2008). Moreover, due to path-dependency—because they breast-feed and need post-birth recovery leaves, and hence spend more time with infants—women tend to be more knowledgeable and competent with, and possibly more attached to, their young children than men. A mother's higher level of competence and attachment does not mean that it is hers, more than a father's, duty to opt out of (full) employment in such a case—the child's interest may be sufficiently well served by the fathers' hands-on care—but it contributes to the psychological pressure on mothers to be the ones who stay at home. And so, a woman's decision to exit the labour market in order to live up to her caring duties is restricted by the lack of acceptable alternatives; when this is the case, the choice is less than fully voluntary.[8] As Anne Alstott noticed (2004), because parenthood comes with a "no exit" clause—parents acquire the duty to provide continuity in care to their children—it significantly restricts parents' autonomy once they have settled into the parental role. And it particularly restricts the autonomy of the parents who are the only hands-on care providers for their children.

The more general point here is that gender norms limit women's autonomy by imposing higher opportunity costs on their, than on men's, pursuit of parenthood (see also Mason, 2000): Combining parenthood with full-time employment is harder for women, who are socially expected to assume the caring role, who tend to earn less them men and who (subsequently) have more difficulties than men in finding a partner willing and capable to assume the role of main care-giver or, at least, to share this role equally. Policies that add to the incentive structure which keeps women away from the market, then, are not necessarily enhancing women's autonomy.

NOT ALWAYS *THEIR* CHOICES

There is an even more important rejoinder to the claim that the negative effects of care-supporting policies on women are morally innocent: Such policies can lower the expected labour market opportunities of one group of women as a result of choices made by another group of women. Here, too, there are several complementary mechanisms to be considered. Women are typically associated with higher social skills, like being caring and competent to meet others' needs, but not with competence in leadership and innovation skills. The retreat from the labour market of some women

---

8.   Assuming that a choice is voluntary if and only if it is not motivated by a lack of acceptable alternatives. See Serena Olsaretti, *Liberty, Desert and the Market* (2004).

can reinforce these biases and thus reinforce explicit sexist beliefs against women in general. More controversially, it can also reinforce implicit biases—that is, a tendency to unconsciously associate certain characteristics with femininity and others with masculinity and to express approval when individuals' behaviour displays the associations attached to their sex and disapproval when it does not (Brownstein, 2015). Scholars of implicit bias claim that virtually everybody harbours such biases, with the result that women are evaluated less favourably in the workplace than men who have the exact same qualifications and accomplishments (Steinpreis, Anders and Ritzke, 1999). Implicit bias may partly explain women's lesser presence in the labour market in general and in some fields in particular, as well as women's difficulties in breaking through various glass ceilings. Biases against women's leadership qualities are much more likely to affect women's opportunities to the more attractive positions. This is because, in existing societies, highly competitive individuals with leadership potential are believed to be the most qualified for the best rewarded positions.

If, as a consequence of a care-supporting policy, enough women retreat from the labour market in general, and possibly from some fields in particular, this will reinforce mechanisms that perpetuate *general* discrimination and marginalisation of women who apply for positions that require leadership abilities and high competitiveness. Hence, women *other* than those who respond to the policy will be even more likely to have their opportunities for positions of advantage curtailed by gender biases.

Further, there exists a fully conscious and rational mechanism that can impose on some women the price of other women's choices to decrease their participation in the labour market: statistical discrimination. If women on average are more likely than men to quit their job or to reduce their labour participation when they have children, and if employee turnover is costly, then it is economically rational for employers to give preference to prospective male workers. Even nonsexist employers who are immune to implicit biases—assuming they exist—will respond with higher levels of statistical discrimination to women's lower participation in the labour market (Bergmann, 1998). Take, for example, the case of hiring scientists who need expensive equipment tailored for the purposes of their research agenda. Each time a lab loses— permanently or for extended periods of time—such an employee they incur a great financial loss (some research projects become soon obsolete so they cannot be simply be put on hold during one's care leave). With these kinds of jobs there are powerful

economic incentives to avoid hiring individuals who are more likely to take significant leaves.

A policy that enables women to avoid the labour market therefore will also, indirectly, cause *other* women to pay for these choices with a more limited set of opportunities for positions in which employee turnover is costly. Thus, as in the case of implicit bias, statistical discrimination is more likely to operate against a woman the better the job to which she aspires: In existing societies such positions usually require high qualifications which, in turn, depend on a commitment to full-time careers. In contrast, the jobs where a high level of staff turnover is not particularly problematic for employers, because it is easy and cheap to train new workers, come with significantly lesser benefits.

Therefore, feminists rightly worry that policies that directly or indirectly protect care-giving will strengthen gender norms against women's access to positions of advantage. These norms are bad for gender fairness, because they make women's and men's access to positions of advantage dependent, in a morally arbitrary way, on them being women or men. This would be serious criticism in circumstances characterised by background distributive justice, where the inequalities in the pay-offs associated with different jobs were just and in which the unemployed—or, at least, the involuntarily unemployed—were not at risk of undue poverty and social marginalisation. Yet, it has little, if any, weight in a world displaying unjustly large inequalities. The rest of the paper explains why.

EQUALITY OF OPPORTUNITIES TO POSITIONS OF ADVANTAGE IN UNJUST CIRCUMSTANCES

When it comes to the distribution of jobs, there are two particularly popular versions of the principle of equality of opportunities (henceforth EO), both of which are meritocratic. One version requires that access to positions of advantage depend only on individuals' talents and ambition. This is often called "formal EO», and prohibits sex-based discrimination. In a more egalitarian, and demanding, version, usually referred to as "fair equality of opportunity" (henceforth FEO), the principle states that individuals' prospects should only be determined by *native* talent and ambition (Rawls, 2001, p.44). In this form, the principle also requires that girls' and boy's talents and ambitions be equally cultivated.

One feminist virtue of both versions is expressive: by prohibiting employers to

give less consideration to an applicant merely on account of their sex, EO principles send the right message that women and men are moral equals. As I argue below, this expressive value can be realised without ensuring that all individuals have equal opportunities for positions of advantage.

The most convincing defense of the more egalitarian version of the principle, usually referred to as "fair equality of opportunity" (henceforth FEO)—is prioritarian. Rawls (2001), for instance, thought FEO regulates competition for social positions in societies that only tolerate those inequalities of outcome that are necessary to incentivise talented individuals to be socially productive.[9] In such circumstances, FEO contributes to the maximisation of the social output which is then distributed according to prioritarian principles; this process maximises the prospects of the worst off. Yet, this defense of FEO depends on prioritarian redistribution actually taking place, and therefore implies nothing about the truth or weight of FEO in societies in such redistribution does not actually take place. Here I assume that existing and predicted levels of inequality in current societies fail to meet prioritarian distributive standards, or indeed any other standard likely to be supported by a purportedly egalitarian theory. They even fail to meet the more modest, sufficientarian standard, that requires that everybody has unconditional access to enough resources to lead a decent life.

We therefore have reasons to doubt that FEO has the same weight in unjustly inegalitarian societies as it has in perfectly just ones—or, indeed, that it is the correct principle regulating competition for positions of advantage. A full analysis of this matter is outside the scope of the present paper.[10] To settle the question at hand, it is enough to examine the likely implications of applying EO to those kinds of positions to which women's access is most likely to be negatively affected by strengthening the mechanisms that already obstruct gender equality. By "positions" here I mean roles with a set of rights, powers and prerogatives of office; thus, salary and political influence are not contingent add-ons to a position, but essential parts of it.

First, the very top positions come with amounts of economic reward, social

9.   For a recent and more explicit defense of this view, see Thomas Scanlon, *Why does Inequality Matter?* (2018).

10.   For more general criticism of different versions of the principle of equal opportunities in societies that fail to observe any distributive principle see John Baker, *Arguing for Equality* (1987, p.46) and Anca Gheaus, "Fair equality of opportunity in unjust circumstances" (manuscript). For an argument that competitions whose pay-offs are very unequal are not made legitimate by ensuring that competitors enjoy equal opportunities see Colin Macleod "Equality of Opportunity and the Consequences of Choice" (manuscript).

power and status that nobody would enjoy in a just society. Some of the positions that lie above the glass ceiling which women find difficult to break are, plausibly, of this kind. The first argument applies to them:

<div align="center">ARGUMENT 1</div>

- P1 Nobody has a right to a position that confers unjust advantages on those who hold it.

- P2 We ought to discount the value of (fair) equality of opportunity for positions to which nobody has a right.

- P3 Many of the top positions in societies with large unjustifiable inequalities of outcome confer unjust advantages on those who hold them.

- C1 Therefore, we ought to discount the value of (fair) equality of opportunity for many of the top positions in societies with large unjustifiable inequalities of outcome.

If C1 is true, improving FEO with respect to top positions cannot be an important goal of justice—and surely not one that is as important as improving the situation of women who are amongst the worse off due to their care role. I assume P1 is self-evident, and so I concentrate on defending P2. One reason why a principle of equal opportunity between women and men is so widely endorsed is that it is obviously unfair to deprive someone of opportunities to a good for morally irrelevant reasons such as one's sex. Yet, this reason is only relevant in cases in which justice permits people to have some opportunity to achieve that good—that is, where it is morally permissible for *someone* to achieve that good. It is unclear that it is morally important to run meritocratically competitions for positions that nobody ought to occupy: nobody has a right to such a position. Ideally, such positions would be abolished; but, if and as long as such positions are up for distribution, a fairer way to distribute them seems to be by lotteries with a threshold of competence (to avoid hiring people that would inflict damage) and that give everyone an opportunity.

There seems to be something very prejudicial to be denied, specifically due to one's sex, the same opportunities that others have even for goods that no one should have in the first place: As already noted, a strong reason to uphold equality of opportunity between women and men is expressive: If a person is denied, merely on

grounds that she is a woman, an opportunity to a good G that nobody ought to have, she is given the message that she is worth less than a man. This gives grounds of complaint not only to her but, plausibly, to *all* women, since the expressive disvalue of discrimination affects women in general. This does not, however, necessarily mean that the woman in question has any claim of justice to more opportunities than she already has to G; if the very existence of G is unjust, she—like everybody else—should have zero opportunities for it. More plausibly, the woman in this case has a special claim (alongside the claim that we all have that G should be abolished) that any opportunities to G be also denied to her male competitors.

Perhaps a meritocratic organisation of competition for goods which nobody ought to enjoy has some unique value, which I have failed to consider here. (The expressive value is not unique to the meritocratic distribution, since this value could also be realised by lotteries or, as I explain presently, a quota system.) In this case, even in unjust societies, ideal feminist policies would include measures that offset the disruption of equality of opportunities between women and men for top positions. But, whether or not EO for top positions has any weight in unjust societies, the most significant progress towards justice can come from improving even *slightly* the situation of the worst off. This is a powerful argument in favour of care-supporting policies.

Here is a thought experiment to test intuitions about the value of improving gender parity for unjust advantages compared to improving the situation of the worst off. Any conclusion about its relevance to the issue at hand will, of course, depend on how you think about justice—especially economic justice—more generally and, therefore, on exactly how unjust you think are our present circumstances. Consider an ancient slave-holding societies that banned women from holding property and hence also from owning slaves. Assume that such a society contemplated a reform granting women the legal right to hold slaves, reform which wouldn't have altered the number of slaves and slave-owners, but would have eliminated any gender inequality of opportunity to be a slave-owner. The reform seems to have some moral value because it eliminates one form of discrimination against women. But, because it achieves this by giving women opportunities to hold an unjust social position, the reform cannot cater to any individual's claim of justice: Nobody should hold slaves, hence nobody has a claim to an opportunity to hold slaves. Therefore, the reform can possibly improve justice only if the demands of justice go beyond fully satisfying the

rightful claims of all individuals involved.[11] In any case, the real value of the reform seems to be expressive: it sends the message that women and men—even if only of a certain social class!—are moral equals. Yet, the reform may also have one kind negative expressive value, because it sends the message that women and men are equally responsible for injustice. The morally sound way of sending the message of women's equality is to deny men, as well as women, the right to hold slaves—that is, to level down, not up.

Some may conjecture that the imaginary reform and, by analogy, reforms that would unblock women's access to top positions have instrumental value in the form of trickle-down effects. In particular, they may undermine the association between women and powerlessness, and as a result improve the opportunities of women across the entire society. Yet, this consequence is uncertain (will the trickle-down really happen?) and it might be cancelled out by the other, negative, effects of associating powerful women with the wrong of slave-holding. Most relevantly, any desirable trickle-down effects would occur due to improved gender parity—that is, by an increase in equality of *outcome* rather than opportunity. If so, then they could be more directly realised through measures that put some women in prominent positions—such as gender quotas attached to lotteries. Any desirable consequences of having more women in top positions can be achieved through reforms—such as lotteries with quotas—that are much less demanding than achieving EO.

More generally, although it is important to undermine gender norms, not any means of fighting them is legitimate: denying care-supporting policies (assuming they are otherwise desirable in terms of justice) to disadvantaged women, in order to give other women more equal opportunities to the top positions is one of the worst ways to attempt to undermine gender norms. As well as being morally objectionable, such strategy is also likely to be inefficient: *some* norms concerning the gendered division of labour are likely to be strengthened by care-supporting policies (Robeyns, 2000; Gheaus, 2008). But *other* biases against women could be strengthened if an otherwise justified policy was opposed on grounds of equal opportunities between women and men for top positions: associating women with enjoyment of unjust benefits can easily backfire by fueling misogyny. It is safe to say, at the very least, that in unjustifiably unequal societies the first concern of justice is to make sure that existing social

---

11.   As I believe is the case; in Anca Gheaus, "The feasibility constraint on the concept of justice" (2013).

positions become more equal; regulating competition for top positions by EO seems much less important.

In defense of the second premise, the slavery example is meant merely as a heuristic to tease out beliefs about the intrinsic value of non-discrimination in access to unjust advantage. Its point, of course, is not to model an analogy to current societies. Slave-owning is unjust in part because it denies some people their rights to freedom from coercion, and liberals typically give priority to such rights—and their fair distribution—over the distribution of other kinds of goods. One may think that current societies in the most developed parts of the world thrive from the dire exploitation of some individuals whose lack of reasonable alternatives makes their situation, in some respects, akin to that of slaves. Examples include firms that rely on sweatshops employing individuals whose alternative is starvation, or firms paying less then subsistence wages to some workers (who, again, lack alternative) although they can afford to pay exorbitant wages to employees in top jobs. Perhaps some of the positions above the glass ceiling that is so difficult for women to crack actually depend on violations of the same rights against coerced exploitation that are violated slavery. It is far from clear that there is any value in improving equal opportunities to positions on the executive board of a company that relies on sweatshops.

But the argument does not depend on the belief that some of top positions involve rights violations on a par with slavery. It is enough to accept that current societies allow massive maldistributions of material goods that are, in some respects, akin to theft. One could run the thought experiment with other kinds of social injustices, which involve violations of lesser rights, such as the right to property. Imagine a group of individuals who trade in stolen goods and fail to give the women and the men in the group the same opportunity to enjoy the revenues. How valuable would it be to ensure that women and men have equal opportunities to share the loot? It is hard to see how individual thieves could have an entitlement to an equal opportunity to benefit from theft, or why denying such opportunities to women represents an injustice. Indeed, it would be odd if the law was to prosecute thieves not only for theft but also for failing to share the loot according to an EO principle—as it should do, if discrimination against women thieves was a violation of justice.

Even if the second analogy, too, is faulty, because the best positions in rich democratic countries are *not* akin to (benefitting from) theft, one can agree with the more modest, comparative claim defended here. If the distribution of economic goods, and, in particular, the sharply increasing levels of inequality and insecurity that affect

lots of people are unjust, then some of the benefits attached to existing top positions would not exist in a just society. On most egalitarian accounts, in a just society there would be no wages and levels of status and power like those enjoyed by pop and sports stars, many CEOs and executive board members, or top employees in Silicon Valley. Yet, in existing societies, gender norms unfairly disadvantage women from having the same access to such benefits as men. We have much less reason to be concerned with equality of opportunity between women and men for such positions than with improvements in the situation of the worst off.

Some may think that, unlike the case of slavery or theft, it is not in itself morally wrong to enjoy pay-offs like those attached to the best positions available in current societies. Sometimes this belief comes from the intuition that it can be legitimate to hold competitions in which the highest prize is similar to a sport star salary while the lowest reward is less than a subsistence wage. I believe that the pull of this intuition depends on such competitions being optional, which makes people's participation in them voluntary. Yet, current participation in the job market is not voluntary, since most of us depend on jobs for survival.[12] Moreover, the best positions in existing societies come not only with enormous *salaries*, but also with levels of power and status that may render them impermissible even if the competition was entered voluntarily.

Therefore, people who, *merely* due to their gender, have lesser opportunities to top positions, lack a claim to more opportunities for such jobs. If so, then there is little reason to worry about how care-supporting policies set back the opportunities of the better off women relative to such positions. The wrongs that statistical discrimination and sexist biases inflict on women who aspire to top positions do not consist in depriving them of equal opportunities to the material goods, status and power of their male peers. The real wrong of, say, paying a male pop star significantly more than a female pop star is expressive, and so the female star's special claim is that the male star's salary be lowered, not that *hers* be increased.

If a consequence of care-supporting policies was the reinforcement of statistical discrimination, this would negatively impact on the opportunities of many more women than just would-be CEO, executive board members, pop and sports celebrities or top employees in Silicon Valley. Women who aspire to less extravagantly well-compensated occupations such as in law, medicine, economics, academia, research or politics—and, especially, at the higher ranks of these and many other professions—will also be negatively affected. Perhaps the benefits associated with *some*

12.   See, again, Olsaretti (2004) for how lack of acceptable options affects voluntariness.

of these jobs would also exist in a just society. But note that it is likely that only very few positions in our society would exist in an egalitarian society. The rights, powers and prerogatives attached to most currently existing positions may be incompatible with just social relations because of the levels of power and status attached to even middle-class typical jobs.

Let me nevertheless concede that some of these positions of advantage would exist in a just society. The second argument concerns such positions and is about the distribution of opportunities for these positions across the entire population:

<div align="center">ARGUMENT 2</div>

· P4 Most people who have a shot at positions of advantage are likely to already have had more than their fair share of opportunities for such position.

· P5 People lack a claim to more than their fair share of opportunities for positions of advantage.

· C2 Therefore, most people who have a shot at positions of advantage do not have a claim to more opportunities than they already have for such positions.

To defend P4: Because in existing societies most people are excluded from opportunities to positions of advantage, a majority of the women who fail to secure positions of advantage *merely on account of their gender* are likely to have already had more than their fair share of opportunities to such positions.[13] How much more will depend on the relevant conception of equality of opportunity.[14] Consider two possibilities: FEO requires that one's opportunities should be determined by native talent and ambition alone. Yet, existing societies are far from giving the same nurturing to all children; not only gender but also class, race and many other factors contribute to unequal nurturing. Given the great inequalities in the cultivation of individuals' talents, it is likely that many people who would compete for any particular position of advantage in a society that realised FEO cannot even *enter* that particular competition in existing societies. Their talents have not been cultivated enough to allow them to qualify formally. Alternatively, assume, more modestly, that equality of op-

13.   Which is not to deny that they are likely to have had fewer opportunities for such positions than similarly advantaged men.

14.   For an overview of different understandings of equality of opportunity see Andrew Mason, *Levelling the Playing Field The Idea of Equal Opportunity and its Place in Egalitarian Thought* (2006).

portunity is satisfied by having careers open to talents; it is very plausible that factors such as friends, family, attendance of the right educational institutions and other forms of brute luck—rather than mere qualification—determine who is in the run-up for many positions of advantage.[15] Given the great inequalities between individuals' pedigrees, social connections etc., it is likely that many people who would compete for any particular position of advantage in a society that realised careers open to talents lack *any* chance in that particular competition in existing societies. Some will not even enter the competition, for lack of information. In other cases, individuals' social connections, pedigrees etc. are not sufficiently good for them to be given real consideration for the job. If so, the majority of women who fail to obtain a desired job *merely on account of their gender* are likely to have already had more than a fair share of opportunities to that position thanks to a better than average education, social connections etc. Ensuring that there is no *gender* disparity of opportunity for positions of advantage would not result, in the absence of additional and more radical reforms, in the level playing field in the way required by the principle of careers open to talents— and even less so by FEO. Rather, it would mostly level the playing field between the already unfairly advantaged women and men.

Again, this is not to conclude that people who fail, merely on account of their gender, to secure jobs that confer unjust advantages on those who hold them, have no complaints of justice. But their complaint of justice is not a claim for *more* opportunities than they already had—which care-supporting policies may impede. Rather, they have a special claim of expressive justice that they not be given the signal that their interests are less important than men's. This is a claim that the men who are favoured merely on account of gender have fewer opportunities than they currently have. This aim, again, can be achieved in the absence of realising EO—for instance, by adopting measures such as lotteries with thresholds of competence and gender quotas. Moreover, even if care-supporting policies entrench statistical discrimination against women, it is not obvious that, on the whole, their consequences for expressive justice will be negative. Precisely because they would sponsor some women's decision to be care-givers, supporters of such policies believe that they have positive consequences in terms of expressive justice; they signal that care—the traditional "women's work"—is valuable.[16]

The general conclusion from the two arguments is that we should discount the

15.   I argue for this in detail in "Three cheers for the token woman!" (2015).
16.   For the case of BI, see Baker (2008).

value of improving women's opportunities to top positions and of improving most women's opportunities for positions of advantage for which they stand a real chance. At the very least, in societies that seriously fail to realise any egalitarian redistributive ideal these aims cannot come close to outweighing the interests of disadvantaged women in care-supporting policies.

What about middle positions that come with decent, though not above average, pay, benefits and security? Examples may include—depending on the society—office jobs or jobs in nursing and teaching. In some places, some such positions—for instance accountancy or human resources—do not even require applicants to have university education, but, instead, an apprenticeship. German societies, for instance, still have desirable jobs that are based on (extensive) vocational education; therefore, it is more likely that individuals who did not enjoy more than average levels of opportunities throughout their lives may have a real shot at these positions. But the kind of statistical discrimination and biases against women that could be amplified by care-supporting policies are a lot less likely to further disadvantage women who compete for such jobs. First, because such jobs are already fairly feminised, a stronger association between women and care need not cause women to be perceived as less competent teachers or nurses. Second, jobs of this kind are often easier to regulate for employees' protection, especially when the employer is the state (as it is the case, in Europe, with most schools and hospitals, and many office jobs). Third, and perhaps more importantly, when women in middle-grade jobs use care leaves or a basic income to temporarily exit the job market, this is unlikely to do the same harm to business as it is the case with jobs that require higher levels of training and specialisation. There will be little reason to discriminate against women in middle-grade jobs due to their involvement in care work.[17]

The last point also applies to lower-grade positions—those with benefits that are significantly below average. But even assuming that care-supporting policies would reinforce women's association with domesticity, and thus have detrimental effects on equality of opportunity for these positions, it is not clear how much of a worry this is. The women whose interests are most likely to be directly advanced by the introduction of BI or care wages are single mothers (some of the poorest people in developed countries); elderly women who spent much of their adult years caring for children,

17.   Senior positions in many such professions are disproportionately occupied by men, and the explanation may have to do with the time women take out for care work. If such senior positions are above average, then they are covered by Argument 2. But if and when they aren't women whose opportunities for them are negatively affected by care-supporting policies have grounds for complaint.

old parents and sick relatives and who, as a result, have meagre pensions; and, in the case of BI, women migrants with little education and social capital. These are not the women whose opportunities are most likely to be negatively affected by implicit and explicit biases or statistical discrimination. Women from poor backgrounds (who are often also from immigrant backgrounds) are some of the least likely to stand a chance to be employed in positions of advantage or in the higher ranks of any profession. The glass ceiling is not a typical problem for women who are chronically unemployed, or in jobs from which they can be easily fired whether or not they are pregnant or saddled with care responsibilities. For the same reasons, these are also the women whose socialisation into caring roles is least likely to deprive them from the non-material benefits of paid work, since they have very few opportunities for jobs with such benefits in the first place.[18]

Further, a large proportion of the low-grade jobs are excessively repetitive and boring, demanding in terms of time, and often organised in unnecessary and oppressive hierarchies. They seem mostly undesirable and, therefore, better opportunities for them seem less beneficial than a secure, substantial and unconditional monthly income or other types of support for care-givers.

Another concern is the effect of care-supporting policies on *men's* opportunities to engage in care work on a par with women. A fully spelled-out ideal of gender justice requires that men have the same access as women to objectively valuable goods such as raising a family, and the same opportunities to avoid bads like exploitative or meaningless work (Gheaus, 2012). The effect of women's exit from the labour market may be to strengthen those gender norms which make it very hard for men to be full time, and sometimes also part-time, care-givers: There is peer pressure on men to behave less caringly than women, there are also ridiculing stereotypes of men who do "women's work", and men receive less training as potential care-givers.[19] Implicit biases against men and statistical discrimination against their presence on the sites of childrearing and in professions that involve direct access to children (like baby-sitter or nursery worker) can make it difficult for them to gain access to some caring profes-

---

18.   Indeed, BI has historically divided feminists along class lines, with working class women supporting it and economically better off women opposing it. See Toru Yamamori, "A Feminist Way to Unconditional Basic Income: Claimants Unions and Women's Liberation Movements in 1970s Britain" (2014).

19.   Against this possibility, Julieta Elgarte (2008) argues that a basic income would provide due protection to caregivers and double-shifters while also enabling women and men to spend less time in paid employment, which increase men's opportunities to become care-givers in their families. The same applies to other care-supporting policies such as gender-egalitarian parental leaves or flexible working hours for care-givers.

sions. But if, in spite of their limited opportunities in this domain, men are overall better off than women, their interest cannot be more weighty than the interest of the worse off women in care-supporting policies. This being said, there is a remaining concern those men who are among the worse off members of society and who would like to take on caring roles;[20] stronger gender norms can make it even harder for such men to access paid care work. I assume they are a very small minority, and therefore this worry does not constitute a reason to oppose care-supporting policies; but it might constitute an additional reason for states to undermine the gendered division of labour by other means.

Therefore, care-supporting policies are least likely to have net negative effects on the worst off women. Advancing the interests of these categories of women carries significant weight—making the case in favour of care-supporting policies stronger than the case against them. These policies include a (suitably-designed) BI, long subsidised care leaves, or split paychecks. When it is possible to means-test care-supporting policies (obviously not in the case of BI but perhaps in the case of care leaves), doing so might be best suited to advance all the feminist goals. Means-testing would avoid incentivising women to stay away from the labour market, and thus would protect EO at the same time as helping the worse off.[21]

CONCLUSION

To the extent to which the introduction of care-supporting policies would bring systematic and enduring improvement for the worse-off women, the fact that it would also set back equality of opportunity between women and men to positions of advantage does not raise a weighty concern. If women make a majority of the disadvantaged members of society, the rebuttal of the equality of opportunity criticism to these policies can be upheld by egalitarian feminists in virtue of their feminist convictions, as well as in virtue of their egalitarian convictions.

The relevance of the present argument goes beyond its implications concerning the introduction of care-supporting policies. An influential strand in contemporary feminism is driven by the ideal of equal opportunities between women and men for positions of advantage. Yet, this ideal has very limited, if any, weight, unless and until two other goals are realised: first, ensuring that the pay-offs associated with positions

20. I am grateful to Thomas Douglas for drawing my attention to this.
21. I am thankful to an anonymous referee of this journal for drawing my attention to this point.

of advantage are permissible, by bringing inequalities of outcome in line with the requirements of justice. The second goal is to ensure that individuals enjoy equality of opportunity at all stages of their lives.

REFERENCES

Anita Allen, "Rationalising Oppression", *Journal of Power* 1(1):51-65, 2008.

Anne Alstott, *No Exit: What Parents Owe Their Children and What Society Owes Parents*, Oxford: Oxford University Press, 2004.

Harriet Baber, "Two Models of Preferential Treatment for Working Mothers", *Public Affairs Quarterly* 4(4):323-334, 1990.

John Baker, *Arguing for Equality*, London: Verso, 1987.

———, "All Things Considered, Should Feminists Embrace Basic Income?", *Basic Income Studies* 3(3):1-7, 2008.

Barbara Bergmann, "The Only Ticket to Equality: Total Androgyny, Male Style", *Journal of Contemporary Legal Issues* 9(75):76-86, 1998.

———, "Long Leaves, Child Well-Being, and Gender Equality", *Politics and Society* 36(3):350-359, 2008.

Michael Brownstein, "Implicit Bias", Stanford Encyclopedia of Philosophy, 2015. Xx partial ref?

"Gender statistics", Eurostat, February 2017. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Gender_statistics_Table1.PNG&oldid=328445 [Accessed 18 June 2020].

"Gender balance on corporate boards", European Commission, July 2016.

Julieta Elgarte, "Basic Income and the Gendered Division of Labour", *Basic Income Studies* 3(3), 2008.

Anca Gheaus, "Fair equality of opportunity in unjust circumstances", manuscript.

———, "Basic Income, Gender Justice and the Costs of Gender-symmetrical Lifestyles", *Basic Income Studies* 3(3):1-7, 2008.

———, "Gender Justice", *Journal of Ethics and Social Philosophy* 6(2), 2012.

———, "The feasibility constraint on the concept of justice", *Philosophical Quarterly* 63(252):445-464, 2013.

———, "Three cheers for the token woman!", *Journal of Applied Philosophy* 32(2):163-176, 2015.

———, "Hikers in Flip-Flops. Luck Egalitarianism, Democratic Equality and the Distribuenda of Justice", *Journal of Applied Philosophy* 35(1):54-69, 2018.

Anca Gheaus and Lisa Herzog, "The Goods of Work (Other than Money!)", *Journal of Social Philosophy* 47(1):70-89, 2016.

Janet Gornik and Marcia Meyers, *Families That Work. Policies for Reconciling Parenthood and Employment,* New York: Russell Sage Foundation, 2003.

Colin Macleod, "Equality of Opportunity and the Consequences of Choice", (manuscript).

Andrew Mason, "Equality, Personal Responsibility, and Gender Socialisation", *Proceedings of the Aristotelian Society* 100(3):227-246, 2000.

———, *Levelling the Playing Field : The Idea of Equal Opportunity and its Place in Egalitarian Thought*, New York: Oxford University Press, 2006.

Claire Cain Miller, "When Family-Friendly Policies Backfire", *New York Times*, 26 May 2016. https://www.nytimes.com/2015/05/26/upshot/when-family-friendly-policies-backfire.html [Accessed 18 June 2020].

Susan Moller Okin, *Justice, Gender and the Family*, New York: Basic Books, 1989

Serena Olsaretti, *Liberty, Desert and the Market*, Cambridge: Cambridge University Press, 2004.

Ann Orloff, "Why Basic Income does not Promote Gender Equality", in Karl Widerquist, Jose Noguera, Yannick Vanderborght and Juregen De Wispelaere (eds.), *Basic Income: An Anthology of Contemporary Research*, Chichester: Wiley-Blackwell, 2013, pp. 149-152.

Janet Radcliffe-Richards, "Only X%: The Problem of Sex Equality", *Journal of Practical Ethics* 2(1):44-67, 2014.

John Rawls, *Justice as Fairness: A Restatement*, Cambridge, Massachusetts: Harvard University Press, 2001. XX place

Ingrid Robeyns, "Hush Money or Emancipation Fee? A Gender Analysis of Basic Income". In R. van der Veen and L. Groot (eds.), *Basic Income on the Agenda. Policy Objectives and Political Chances*, Amsterdam: Amsterdam University Press, 2000, pp. 121-136.

———, "Will a BI do justice to women?", *Analyse und Kritik* 23:88-105, 2001

Thomas Scanlon, *Why does Inequality Matter?*, Oxford: Oxford University Press, 2018. XX Place

Gina Schouten, *Liberalism, Neutrality, and the Gendered Division of Labor*, Oxford: Oxford University Press, 2019.

Rhea E. Steinpreis, Katie A. Anders and Dawn Ritzke, "The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study", *Sex Roles* 41(7):509-528, 1999

Claude M. Steele and Joshua AronsoN, "Stereotype threat and the intellectual test performance of African-Americans", *Journal of Personality and Social Psychology*, 69:797-811, 1995.

Gijsbert Stoet and David C. Geary. "Can stereotype threat explain the gender gap in mathematics performance and achievement?", *Review of General Psychology* 16:93—102, 2012.

Lawrence J. Stricker and William C. Ward, "Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance," *Journal of Applied Social Psychology* 34:665-693, 2004.

Yannick Vanderborght and Karl Widerquist, "The Feminist Response to Basic Income", in K. Widerquist, J. Noguera, Y. Vanderborght and J. De Wispelaere (eds.), *Basic Income: An Anthology of Contemporary Research*, Chichester: Wiley-Blackwell, 2013, pp. 142-144.

Ann Whithorn, "Is One Man's Ceiling Another Woman's Floor?", in K. Widerquist, J. Noguera, Y. Vanderborght and J. De Wispelaere (eds.), *Basic Income: An Anthology of Contemporary Research*, Chichester: Wiley-Blackwell, 2013, pp. 145-148.

Joan Williams, *Unbending Gender: Why Family and Work Conflict and What to Do About It*, New York: Oxford University Press, 2000.

Toru Yamamori, "A Feminist Way to Unconditional Basic Income: Claimants Unions and Women's Liberation Movements in 1970s Britain", *Basic Income Studies* 9(1-2):1-24, 2014.

Duty and Doubt

SETH LAZAR

Australian National University

ABSTRACT

Deontologists have been slow to address decision-making under risk and un-certainty, no doubt because the standard approaches to non-moral decision theory appear superficially similar to consequentialist moral reasoning. I identify some central tenets of simple decision theory and show that they should not put deon-tologists off, before showing where we should go next to develop a comprehensive deontological decision theory.

1. INTRODUCTION

As any bad gardener knows, the symptoms of overwatering and underwatering are often the same. The leaves start to yellow, the stems lose stiffness and vigour. If you do nothing, the plant dies. But if you base your response on the wrong diagnosis, you actively accelerate that process. Sometimes you can do more to discover what is wrong before acting, but the evidence is often inconclusive.

Almost all human decision-making takes place under uncertainty. When the stakes are purely prudential—whether my camellia flourishes or dies—this is chal-lenging enough. But moral decision-making is no less dependent on facts. And the facts are just as commonly in doubt.

Doctors, of course, know this. The injunction to 'first do no harm' is oblique at best. It can only plausibly be interpreted as 'first, do no harm, all things considered, in light of the information you ought to have'. In countries vulnerable to wildfire, firefighters know this too. Many of Australia's fiercest bushfires have resulted from 'hazard reduction' burns, intended to reduce danger to life and property, but which

would never have been started had those in charge known how the weather conditions would change.

And soldiers at war know it. Indeed, 'the fog of war' is so endemic as to be a cliché. Life and death decisions must be made at all junctures—often without knowing who will live and die, whether they are combatants or noncombatants, and indeed precisely what one is fighting for.

Although this combination of high stakes and high uncertainty is normally the province of emergency practitioners, those with more pedestrian lives also face risky decisions, albeit usually with less uncertainty. For example, many of us routinely operate lethal projectiles at high speeds, knowingly imposing a low risk of substantial harm on those we drive past, often for the sake of trivial gains (a trip to the shops for chocolate, perhaps). Some of us eat meat, running the risk of grievously wronging animals that might in fact enjoy moral status. We try to raise our children to be fulfilled and secure—for many of us no other obligation weighs heavier. But we are forever making consequential decisions with incomplete information, which might, instead, lead to the result that Philip Larkin predicted.

Practically all real-life moral reasoning takes place under uncertainty. That's not news. And yet a sizeable wing of moral philosophy has failed to adequately acknowledge this. What's more, it's the wing that otherwise seems closest to the truth.

There are many different ways to understand 'deontological' ethics. The simplest is by defining the opposition: 'consequentialism'. Consequentialists think, very roughly, that when choosing what we are morally permitted to do, we must consider the outcome that each available act will realise. To each outcome we assign a value; we are morally permitted to perform an act if, and only if, no other act has a better outcome.

By contrast, deontologists think that sometimes it is permissible to perform an act, even though an alternative would have a better outcome. In other words, sometimes doing less than the best is not morally wrong. This is most obviously true when the act with the better outcome involves excessive personal sacrifice, in relation to the good done.

Suppose I have the last ticket for the last performance of a play you and I both want to see. And suppose that you would enjoy it somewhat more than I would. Hold everything else constant. The outcome where I give you the ticket is somewhat better than the one where I view the play myself: pleasure is good, and you'll get more pleasure from the play than I will. By consequentialist lights, I am morally required to

give you the ticket. Deontologists will baulk at that result. We could all make the world a much better place, if we had no regard for our own well-being. It might be heroic, or saintly, to sacrifice ourselves for the sake of marginally greater benefits for others. But doing so is not, in general, morally required.

Deontologists also think that sometimes it is wrong to bring about the best outcome (on a natural understanding of 'best'). Suppose that you could ensure that more children were lovingly cared for by their parents, by neglecting your own children. Or that you can minimise promise-breaking by others by breaking your promise now. Or that if you kill one innocent person, you can use their organs to save five other innocent people's lives. In each case, deontologists think that you are morally required not to bring about the outcome that is (apparently) best. It's not okay to minimise breaches of a given duty overall; you must attend to your own wrongdoing first and foremost.

I'll say more about the deontology/consequentialism divide in the next section. But here is the key difference. Consequentialism is in one sense more demanding than deontology—there is no scope to prioritise your own interests. But it is also much more permissive: it licenses imposing costs on some if you can thereby realise a slightly larger benefit for others.

There are many deontologists in the world (I am among them). But the other side has, as is often remarked, one apparent advantage. Consequentialism pairs very nicely with standard 'decision theory'—the theory of how to rationally make decisions with imperfect information. Deontologists, on the other hand, have offered no systematic account of how to handle decision-making under doubt.[1] That's my project: to develop a theory of duty under doubt or, in other words, a *deontological decision theory*.

2. WHAT IS DEONTOLOGY?

It will help, first, to flesh out what the 'deontological' in deontological decision theory means. Some philosophers have offered subtle structural diagnoses of this 'fault line in ethical theory' (e.g. Nair 2014). But for my purposes here, I'm less inter-

---

1.   Colyvan et al. (2010); Olsen (2018); Spector (2016); Zimmerman (2008) all fail to accommodate supererogation, options to act suboptimally, and in general the justifying/requiring distinction. Nor do they offer an argument for *why* deontologists should adopt expected-utility theory. Aboodi et al. (2008); Zollman et al. (2017) are intentionally very limited in scope. Even its author does not endorse Isaacs (2014). My own approach in this paper draws on work including Lazar (2017a, 2017b, 2018a, 2018b); Lazar and Lee-Stronach (2018).

ested in structural features of ethical theories, and more concerned with their substantive commitments. Here, then, is a simple deontological credo.

- beings with sufficiently advanced rational and moral capacities enjoy moral status;

- beings with moral status have fundamental rights, corresponding to duties that others have not to harm them, and to aid them;

- our duties not to harm others are more demanding than our duties to aid others (or to advance better outcomes in general);

- we have special duties to care for and help our loved ones and friends;

- the manner in which one's action harms or fails to aid someone matters for its permissibility, and our reasons, intentions and beliefs can likewise matter;

- we should be sceptical about justifications for the few bearing significant costs, which appeal to individually smaller, but collectively greater, benefits to others;

- our actions can wrong others, even when we act permissibly, and this 'pro tanto' wronging demands a response (perhaps apology, perhaps reparation);

This grab-bag of commitments is obviously not shared by all deontologists. What's more, plenty of consequentialists would endorse at least some of them. Nonetheless, consequentialists are much more likely to deny these claims, and deontologists to affirm them. My project is to develop a decision theory for *this* kind of deontology.

3. CRITERIA OF SUBJECTIVE PERMISSIBILITY

Except in special cases, decision theory does not aim to provide an operationalisable decision procedure. It aims, instead, to develop a criterion of subjective permissibility. What does that mean, and why should we want one?

We are interested in *moral* permissibility, and *moral* wrongness. (I use 'wrong' and 'impermissible' interchangeably.) Objective permissibility is permissibility in light of all the non-moral facts that could bear on your action. Think back to the

hazard reduction burn. If in fact the weather will change this afternoon, blowing a hot wind towards town, then it is objectively wrong to begin the burn now.

Subjective permissibility, by contrast, is permissibility in light of your imperfect information. This is an intentionally generic formulation, which covers a range of different ways in which your information could be imperfect. We might care, for instance, about whether an act is permissible in light of your beliefs about the world. Or we might care most about your evidence. There are other possibilities besides. In this paper, I will be neutral between them, focusing on questions that apply whichever subjective epistemic standard is at stake. Return to the firefighter. Suppose that, in the morning, there was no reason to believe that the wind would change in the afternoon. Then it may have been subjectively permissible to start the hazard reduction burn. Conversely, if a reliable forecast predicted the change, starting the burn would probably be subjectively impermissible.

A *criterion* of subjective permissibility is a set of necessary and sufficient conditions for an act being subjectively permissible. Ideally, these conditions would not only pick out subjectively permissible acts, it would explain why they are subjectively permissible. We're looking for conditions that state an act is subjectively permissible if and only if, *and because*, X, Y or Z.

This leads to the first substantive challenge to my project. Why should we want a criterion of subjective permissibility? One can understand this objection in different ways.

First, why bother with uncertainty at all? Why not let moral philosophers focus on decision-making under certainty, and leave decision-making under uncertainty to other branches of philosophy, indeed to other disciplines? Aren't these really technical questions about implementation, for which technical disciplines are necessary?

This objection might make sense if we were wholly clueless about subjective permissibility, lacking any intuitive insight into it. Assuming that we are more confident about our judgements of objective permissibility, we might then agree to outsource responsibility for decision-making under risk and uncertainty to a more technical discipline, anchored in our objective moral theory.

However, that is emphatically not where we stand. Our judgements of subjective permissibility are often as robust as our judgements of objective permissibility. For example, if the firefighter thinks it is more likely than not that the wind will change that afternoon, then without some compelling countervailing reason, it would be clearly wrong to start the burn in the morning. If we can draw conclusions like these,

then extending moral theory to decision-making under risk is not a purely technical task, and anyone engaged in it will be doing moral philosophy.

Other critics of criteria of subjective permissibility might argue that philosophers should focus on helping people *make* better decisions under uncertainty—developing actual decision procedures—rather than on criteria of right action (e.g. Smith 2018). My own background is in the ethics of war. I see the urgency of developing better decision procedures. But it is hubristic for philosophers to claim more than a bit-part in that process. Good decision-making involves much more than what philosophy alone can offer. Most notably, developing good decision procedures requires practical wisdom and experience—qualities that philosophers are not widely celebrated for possessing. Of course, we should contribute as best we can—and articulating criteria of subjective permissibility may be one of the most useful things we can do.

I agree, of course, that moral theory should be *action-guiding*. But this does not mean providing a formula that people can apply to actual decisions in realistic circumstances. Instead, it means *ranking* the options available to you, and saying which are permissible, in light of how they appeared to you at the time.

Others might reject 'deontological decision theory' just because it reduces the complex moral landscape to a few simple axioms. Deontologists are normally keen on drawing fine distinctions, and articulating complicated domain-specific principles. They may understandably feel uncomfortable with criteria of subjective permissibility, and instead prefer to accommodate risk in different domains of human activity with different principles (e.g. Bolinger 2017; Ferzan 2005; Quong 2015).

No doubt there is something to this idea. These domain-specific approaches, like deontological decision theory itself, are in their infancy. Both approaches should be developed, so that we can properly decide between them. At a first pass, though, I am sceptical about the domain-specific approach. However we divide up the relevant domains, we will find cases we cannot securely allocate to a specific domain. What principles do we apply then?

For example, suppose one set of principles governs beneficence under risk, while another set governs harming—that seems very likely. What do we do, then, if we don't know whether our options will benefit or harm? Which domain-specific principles do we apply? I suspect that solving this problem in a principled way will lead inexorably to adopting a criterion of subjective permissibility.

A final observation: computer scientists and engineers are already building artificial agents to make consequential decisions. Artificial intelligence is fundamentally

grounded in probability theory and decision theory. Some of the reasons for scepticism that a human agent could ever actually deploy a criterion of subjective permissibility do not apply to artificial agents. They can gather and process much more information than humans can, at least within a narrow domain of expertise. They lack emotions, so cannot be blinded by fear or panic. A further reason to develop a deontological decision theory is that, if we don't, the AI agents of the future will all be consequentialists.

4. THE PROMISE AND PERILS OF
ORTHODOX DECISION THEORY

Perhaps deontologists have ignored decision-making under uncertainty because a separate, interdisciplinary field of enquiry has already claimed that territory as its own. Economists, computer scientists, engineers, mathematicians, and others all contribute to, and make use of, various forms of decision theory. One might think that a group as smart as that would solve any problems that moral decision-making under risk might raise. And indeed deontologists have much to learn from decision theory. But there are some hurdles to overcome first.

Decision theory is, in essence, the attempt to devise a criterion of subjective permissibility for *rational* decision-making under risk. Vanilla decision theory says that an act is subjectively (rationally) permissible if and only if, and because, it maximises expected utility.[2] To understand this principle, we need some ingredients.

First, *acts*. An act is just something that it is in your control to do. It's easy to problematise this notion, but we'll work with a simple, intuitive sense of your available acts. For example, start the burn, or don't; fire at the person in your sights, or don't; and so on.

Second, *states*. A state is a complete description of the world (though to save time we can simply focus on the salient bits). Importantly, it is not just a description of how the world is at a time. It includes, at least in principle, the full history of the world too.

There are many ways the world might be. We do not know which one is actual. But some are more likely to be actual than others. *Probabilities* represent this idea.

2.   Of course, this is as disputatious a field as any other in philosophy, so this characterisation of orthodox decision theory might itself be challenged. For two helpful overviews, see Briggs (2014); Buchak (2016a).

They raise many philosophical and mathematical questions, but we must ignore those here. They provide a measure from 0 to 1 of how likely a state is to be actual. On one appealing formulation: 1 is logically necessary (like a tautology); 0 is logically impossible (like a contradiction).

Economists often define acts as functions from states to outcomes. An outcome is what happens if you perform a specific act, given that a particular state of the world is the case. Again, outcomes are full descriptions of the world, including its history—in particular, including the fact that you performed that act.

To decide between acts, orthodox decision theory has us compare the outcomes that they might realise. We therefore need some way to represent which outcomes are better and worse than others. *Utilities* are the numbers that we assign to outcomes, to measure their level of normative support. The use of numbers here might make some nervous—numbers have properties that reasons might lack, such as precision and simple additivity. The numbers, however, are just a way of comparing the different possible outcomes. Those comparisons are what matters, not the specific unit or scale in which we are measuring. And the use of numbers to represent a comparison is consistent with denying that our reasons have other properties associated with numbers. Some might also be put off by the word 'utility', invoking as it does the paradigmatic consequentialist theory, utilitarianism. That would be premature. The concept of 'utility' is just a measure of normative support. As we will see below, it is flexible enough to accommodate non-consequentialist theories of what matters.

With these ingredients in place, we can define the *expected utility* of our available acts. Consider the decision whether to start the hazard reduction burn. To calculate the expected utility of doing so, we have to determine the possible outcomes of this choice, and how likely they are. To keep things simple, let's imagine that there are only two relevant states: either the weather conditions will remain as they are, or else they will change unfavourably. If they remain as they are, the burn will proceed without a hitch. If they change unfavourably, then a catastrophic fire will ensue. To calculate the expected utility of undertaking the burn, we must assign utilities to these two outcomes, and then multiply each by the probability of the associated state being the case. We then sum the products to get the expected utility, which is a probability-weighted average of the utilities of the possible outcomes of your act. Orthodox decision theory says that it is rationally permissible to start the fire only if no alternative has greater expected utility.

This simplistic overview tells you how orthodox decision theory selects the per-

missible option from a risky choice. One question for deontologists, then, is whether they can adapt this approach to accommodate their view of what matters. However, there is a more pressing concern. Much—indeed most—work in decision theory focuses not on a criterion of right action, but on what justifies it. And the most commonly cited justifications seem ill-suited to a deontological approach to ethics. Before we ask whether deontologists *can* adopt and adapt orthodox decision theory, we must ask if they should want to.

There are three popular arguments for maximising expected utility. They are mathematical, and complex, so I will only sketch them here.

First, the most obvious and commonplace argument for maximising expected utility is that, if you do so, then in the long run and on average, you will realise more *actual* utility. Maximising expected utility in the short run is a great way to get what you want in the long run.

Second, many decision theorists think that one ought to maximise expected utility because any agent whose preferences over gambles (that is, acts with uncertain outcomes) obey some seemingly innocent 'axioms of rationality' can be mathematically *represented* as an expected-utility maximiser. The normative justification for their decision rule, then, is the plausibility of these underlying axioms. Indeed, many decision theorists would think it unbearably jejune even to suppose that they aim to defend a criterion of right action, and in particular to imagine that we can ever identify which act to choose under risk by assigning probabilities to states and utilities to outcomes. They argue that one's preferences over gambles are the only real psychological phenomenon—probabilities and utilities are a construct generated by the 'representation theorem' (we can call these people 'constructivists'). They would argue that orthodox decision theory gives no guidance on how to choose between risky options—only requiring that you obey the axioms of rationality.

Third, Graham Oddie and Peter Milne argue, roughly, that an additional 0.01 probability of some outcome occurring should have the same bearing on your decision regardless of whether the probability of that outcome has gone from 0.9 to 0.91, or from 0.1 to 0.11 (Oddie and Milne 1991). They then show that, if this premise is correct, we are committed to expected-utility maximisation (the details don't matter here).

Notwithstanding the scepticism of constructivists, orthodox decision theory is more than just an elegant bit of maths. It underpins decision-making in many spheres of human—and non-human—endeavour, from governments to insurance compa-

nies, banks to casinos, self-driving vehicles to stock-trading algorithms. It is core to many different academic disciplines, from the behavioural sciences to some branches of physics. It has well earned its title as orthodoxy.

At first sight, this looks like bad news for deontologists. Let's start with the justifications. The first is the most obviously problematic. Justifying one's decisions now because of the long-run implications of deciding in that way is emphatically *not* the deontological approach! We are bound by constraints, which means that sometimes the ends *don't* justify the means. What's more, we've also seen that the ends don't necessarily *require* the means. If you are permitted to act suboptimally in individual choices, then there is no guarantee that you will realise the best outcomes in the long run. We must start by binning the justification grounded in the long run.[3]

Alas, the representation theorem justification is hardly more promising. First, constructivist decision theory clearly fails to be action-guiding. It cannot tell you which of your options maximises expected utility until you have made your choice. It can do no more than rule out options that would violate an axiom of rationality. But *moral* decision theory will need to be more robust than that. The axioms of rationality are merely formal constraints, demanding, in essence, nothing more than consistency. It is perfectly possible to be a consistent moral monster. Constructivist decision theorists harrumphing about probabilities and utilities not being real have abandoned normative theory altogether. Decision theory can—and must—do more than tell us how to be consistent.

Worse still, a deontological approach to ethics sits particularly ill with these supposedly innocent axioms. For example, *transitivity* means, roughly, that for any three options A, B and C, if A is preferred to B, and B to C, then A should be preferred to C. Deontological decision-making under risk is quite likely to violate transitivity. Agents typically have a range of permissible options, and the 'permissible to do X rather than Y' relation is emphatically not transitive, as Frances Kamm has shown. For example, it is (A) permissible to go golfing rather than (B) save a drowning child at high personal risk. It is permissible to (B) save a drowning child at high personal risk rather than (C) keep an appointment. But it is *not* permissible to (A) go golfing rather than (C) keep an appointment (Kamm 1985). Others have shown that the same is true of moral requirements, and indeed perhaps moral goodness (Temkin 2012; Voorhoeve 2014).

If deontologists reject transitivity, then they cannot be represented as expected

3.   As many others have argued, e.g. in this context Tenenbaum (2017).

utility maximisers. They might also plausibly reject *continuity*, which roughly says that you should be prepared to trade off any sure thing against some probability of any other thing. And *completeness* is equally controversial, since it implies that we can, in principle, rank every option in every gamble against every other one—which is impossible if some options are incomparable to others.

Deontologists cannot reverse into orthodox decision theory via a representation theorem. They also can't justify adopting this approach on the grounds of its long-run success. So should they look elsewhere for inspiration? Is orthodox decision theory a dead end?

I think not. Oddie and Milne's argument, above, may point us in the right direction. It focuses attention on just what one is committed to, when one adopts the decision-theoretic framework. If we can pinpoint its fundamental elements, then perhaps we can either argue for those, or show that they cannot plausibly be controverted.

What is left of orthodox decision theory when we jettison long-run justifications and representation theorems? Suppose that we cannot *represent* you as an expected utility maximiser. You can still assign probabilities and utilities to states and outcomes, combine them to determine each act's expected utility, and then use those expected utilities to choose an action. Indeed, if you are going to make defensible decisions under risk, then you must find *some* way to combine a measure of your uncertainty with a measure of what matters. This is simply common sense: any risky choice involves an assessment of both what is at stake and what the odds are. The big question faced by deontologists is whether they should combine these in the way recommended by orthodox decision theory.

In this spirit, I think we can extract two basic ingredients from orthodox decision theory. Representation theorems and the maximising decision rule are not part of the core package. The core package is two simple theses, which I think deontologists can (with a little interpretation) adopt.

First, we can sensibly determine the level of normative support of the possible outcomes of our actions, and rank them accordingly, on one or more dimensions: I'll call this the *Ranking* hypothesis. Second, when deciding under uncertainty, we should discount an outcome's level of normative support (in some dimension) in linear proportion to its probability of coming about: I'll call this thesis *Linear Discounting*.[4]

For the project of developing a deontological decision theory to get off the

---

4.   This is, I think, the same as Smith's 'Evaluative Compositionality': Smith (2014). See also Tenenbaum's description of the 'multiplicative model': Tenenbaum (2017).

ground, we need only to argue that *Ranking* and *Linear Discounting* are consistent with deontological ethics, and that we have positive reason to adopt them rather than some alternative.[5]

4.1 RANKING

Deontologists should not be deterred by *Ranking*. It requires only that, when figuring out whether to perform some act, the reasons for and against it, *within a particular dimension*, are roughly comparable to one another, so that we can somehow aggregate them. That is all. It does not presuppose that we can always *precisely* compare our reasons. Most importantly, it does not presuppose that our reasons all operate within the same normative dimension or domain. Vanilla decision theory's decision rule *does* presuppose that all reasons weigh on a single dimension. This explains why, according to the orthodoxy, it is irrational to choose a worse option when a better is available. But *Ranking* is consistent with endorsing a different decision rule, which registers different dimensions of normative strength (I'll explain this in more detail below).

Ranking does not assume that reasons interact with one another in a simple additive way. Some people think that if a morally relevant property tells for or against your action to a particular degree in one choice, it must do the same in any choice. On this view, a fact's reason-giving force is much like an object's mass in a controlled environment: if you add two properties together, you get the sum of their normative weights when separate. *Ranking* is consistent with this view, but also with its counterpoint—the idea that morally relevant properties can interact in unusual ways beyond mere addition, for example cancelling one another out or amplifying one another's force (e.g. Dancy 2004).

Ranking is conservative, but not vacuous. It insists that our reasons are comparable. For any two considerations, we can judge that one is weightier than the other, or they are equally weighty or roughly equal. If two considerations are incompara-

5.   Notice that one of the main precursors to this paper, Colyvan *et al.* (2010), doesn't offer any argument to justify adopting orthodox decision theory and, in particular, does not defend either *Ranking* or *Linear Discounting*. Instead, it assumes that orthodox decision theory is the way to go, and takes the main challenge to be working out whether a utility function can adequately represent deontological moral theories. What's more, it does this without attending to the most challenging feature of deontological ethics—its commitment to supererogation and other options to act suboptimally. Oddie and Milne (1991) argue that all deontologists can be represented as expected-utility maximisers. They are wrong, again because of the failure to accommodate options to act suboptimally, which clearly lead to violations of transitivity. As far as I can make out, though Spector (2016) and Olsen (2018) endorse *Linear Discounting*, they offer no argument for it.

ble, none of these relations between them obtains. I think deontologists (and others) should enthusiastically give incomparability the boot. What's more, risky choices make this conclusion all the more compelling.

Here's the argument against incomparability. Suppose that two reasons are incomparable with one another: say, your reason not to kill an innocent person and your reason to preserve the natural beauty of the environment. Suppose that your parliament is about to vote in favour of allowing resource extraction in a national park. Your only way of preventing this from happening is to detonate an explosive, which will trigger a sequence of events leading (let's assume) to the parliament reversing its course and protecting the park. The downside is that if you detonate the explosive, you will definitely kill an innocent person. Simplifying a lot, suppose that the only considerations at stake here are the preservation of the environment's natural beauty and the innocent life that you will be taking. If these considerations are incomparable with one another, then there is simply no way to rationally weigh them. We cannot say that one is more important than the other, or that they are (roughly) equal.

Now this is hard to swallow, just on its face. But notice the further implications of endorsing incomparability. In any actual choice, there will be many other considerations at stake. Even if all of those are fully comparable, the presence on either side of incomparable considerations means that your options will be incomparable. That's already a problem (MacAskill 2013). Worse still, this problem arises even if there is a very low probability that incomparable considerations will be at stake. Suppose, for example, that the bomb is very unlikely to kill an innocent person, but sure to save the national park. One might naturally want to infer that the importance of saving the national park outweighs the risk of killing an innocent person, when that risk is low enough. But if these considerations are genuinely incomparable, then that inference is unwarranted.

Here's why. Let's use N as shorthand for natural beauty, and K as shorthand for the wrongness of killing an innocent person. If N and K are incomparable with one another, then any multiple of N and of K must also be incomparable with one another. If they were not, then we would be able to infer how N and K compare. Suppose, for example, that one wanted to say that, though N and K are incomparable, N is more important than $0.01(K)$. If that were true, then one could proceed by a series of comparisons, asking whether progressively larger probabilities of K are still less important than N, until we reach a point where they are roughly equal. If $p(K)$ is

roughly equal to *N*, then *K* must be at least roughly equal to *N*, if not more important than it. So if p(*K*) is comparable with *N*, then *K* is comparable with *N*.

If you think that some reasons are incomparable under certainty, then you must also believe them incomparable under risk. And if the presence of any incomparable considerations makes it impossible to rationally choose between otherwise comparable options, then whenever there is any risk of a choice involving incomparable considerations, the options will themselves be incomparable. The result would be total paralysis. Incomparability, as Will MacAskill has nicely put it, is infectious (MacAskill 2013). That is not a tolerable or plausible result. Deontologists should reject incomparability, and endorse *Ranking*.

4.2 LINEAR DISCOUNTING

Nothing about discounting our reasons in proportion to their probability of being actual conflicts with core deontological commitments. If we should reject *Linear Discounting*, then, it is not because we are deontologists, but because there are independent reasons to reject it.

One might argue that probabilities are not the only tools for navigating an uncertain world. Most notably, we can distinguish between probabilities on the one hand, and 'all-out beliefs' on the other (Isaacs 2014; Tenenbaum 2017). We find something like this in the criminal law. Juries are instructed to convict if they *believe, beyond a reasonable doubt*, that the suspect is guilty.[6] When things go wrong—like for the firefighter whose hazard reduction fire burns out of control—we often ask whether someone *believed* they were acting objectively permissibly, without regard to probabilities. Perhaps we could extend this kind of model to other cases, and so do without *Linear Discounting*?

I cannot debate the merits of these different approaches in this paper; nor should I have to. The prominent role of probabilities in philosophy, the natural and social sciences, and social life more generally, is indisputable. Since Cicero, people have agreed that 'probability is the very guide to life'. Perhaps other guides are available. But it is at least defensible to base one's decision theory on the guide that has the most currency.

There is a second line of argument against *Linear Discounting*, which we cannot so easily dismiss. One might agree that moral decision-making under doubt should

6.   For a helpful overview, see Gardiner (2017).

use probabilities, but argue that probabilities should discount utilities in a different way. Perhaps we should totally disregard very low-probability states (Smith 2014). Or else perhaps we should give additional weight to very high-probability states, treating them as equivalent to certainty.[7] Or we might give disproportionate weight to crossing the 50:50 threshold (Haque 2012). Alternatively, we might give some weight to 'global' properties of gambles, caring about the distance between the worst-case and best-case outcome, or else giving extra weight in our deliberations to one of those extremes. All of these and others could justify rejecting *Linear Discounting*. What can be said against them?

It is surprisingly hard to justify *Linear Discounting*. Oddie and Milne, as pointed out above, thought it just obviously true. They asked, incredulously, what could possibly make an additional 0.01 probability of some outcome matter more when added to 0.1 than when added to 0.9 probability of the same outcome. But this is really just table-thumping. I'm going to offer two arguments, which I hope will be more convincing.

The first raises a worry about attaching special significance to crossing particular probability thresholds—discounting low-probability outcomes, for example, or giving special weight to the 50:50 threshold. This approach makes the expected utility of an option depend acutely on how you define the relevant states. For example, suppose that the firefighter is permitted to ignore states with a lower than 0.01 probability. Before, I described the relevant states as being 'weather remains the same' or 'weather changes unfavourably'. But we could equally well divide the latter possibility into arbitrarily many sub-possibilities (the wind increases by n km/h and shifts to the west, the temperature increases by n degrees, etc). What stops us dividing up the states until they fall below the threshold of indifference, and can safely be disregarded? If there is no obviously privileged way to partition the relevant states, how do we calculate the expected utility of an option?

My second argument gives a general reason to favour *Linear Discounting* over the alternatives. Our goal is to determine how to discount utilities by their probability of being actual. The product of that discounting is literally a product—the result of multiplying one factor by another. Any product that you can reach by tweaking one of the factors (for example, by altering the probability discount) you could instead reach by tweaking the other (by building it into the utility of the associated outcome). The question, then, is not whether any practical verdicts warrant a departure from

7.   On the idea of a 'moral certainty', see Aboodi *et al.* (2008).

Linear Discounting. Instead, the question is whether those verdicts are best illuminated and explained by rejecting *Linear Discounting*, or else by tweaking the utility function.

Our deontic verdicts—what we are permitted to do in a given choice—are determined by the product of our probabilities and our utilities. We can reach precisely the same set of deontic verdicts with a decision theory that endorses *Linear Discounting* as we can with one that rejects it. The question is: which approach is better motivated and has more explanatory power?

Here is my simple thought: expected utilities combine epistemic and practical reasons. Probabilities are the epistemic component; utilities are the practical component. *Linear Discounting* is the default starting point. Probabilities share out the possibility of being actual in the way mandated by the evidence, or the agent's beliefs. *Linear Discounting* shares out the weight of possible utilities in just the same way. Your reasons in a state get to count in proportion to their possibility of being actual. I think that any departure from *Linear Discounting* should be justified on epistemic grounds. If the motivation for abandoning *Linear Discounting* instead comes down to one's sense of the relevant practical reasons, then it should be incorporated within one's utilities.

Here's an example. I think that when you breach some duties, your act is objectively worse (other things equal) the greater your degree of belief, when you acted, that your action would breach that duty. If the probability of your act breaching that duty was very high when you acted, then your act is objectively worse than if it was very unlikely to lead to that breach. So, when a combatant kills an innocent person in war, that killing is objectively more wrongful the likelier it was, when she acted, that the victim would be innocent (Lazar 2015). The reason: taking a bigger risk of breaching one's duty to another involves, other things equal, showing them a greater degree of disrespect. One *could* incorporate this into one's theory by proposing a departure from *Linear Discounting*. But this is fundamentally a practical reason—having to do with the respect that we owe to others—not an epistemic one. The proper place to figure it into our deliberations is among our practical reasons.

Here is another way to see the same point. Probabilities interact with our objective practical reasons in different ways. For example, our duties of rescue are fundamentally epistemically constrained, in this sense: the duties I have depend in part on the information that I have. There is someone in the world, right now, who will die in the next five minutes. I could save their life by phoning to warn them of the

impending threat. The cost to me would be negligible; the benefit to them would be their life. But I do not have a duty of rescue, because there is no way for me to know whom to call. Conversely, I think our duties not to harm others are not epistemically constrained in the same way. If I kill someone accidentally, and there was no way for me to know that my action would cause their death, I think I have still breached a duty to that person (albeit that I may be fully excused) (Thomson 1986).

If this is right, then any departure from *Linear Discounting* will have to be highly context-specific. It will depend, indeed, on the precise contours of the outcomes realised by one's options. This is in contrast, for example, with Lara Buchak's 'Risk-Weighted Expected Utility' Theory, which assumes that our risk attitudes can be represented by a constant function, which does not vary with context.[8] If the probability discount is being determined by the content of the outcomes anyway, then Occam's razor suggests we should just fold this information into the utility assignment.

Deontologists have no special reason to reject *Ranking* or *Linear Discounting*. Nor are there general reasons to reject either. *Ranking* is a central tenet of practical reasoning generally; without it we would be quite lost; and any apparent challenges to *Linear Discounting* are better catered for by recognising that the objective weight of our moral reasons can vary depending on the information that the agent has when she acts. So deontologists have good reason to adopt vanilla decision theory—if they can.

5. CAN DEONTOLOGISTS USE ORTHODOX DECISION THEORY?

Headline: deontologists can adopt the concept of expected utility, but had better reject the vanilla decision rule. Some will baulk at the first point. Doesn't the concept of expected utility require us to *evaluate outcomes*? Isn't the whole point of deontological ethics that it doesn't go in for that?

Some philosophers certainly take this view. They have been riled by others who have sought to 'consequentialise' moral theories—not only deriving a *ranking* on outcomes from every plausible moral theory, but pairing that *ranking* with a maximising decision rule, so that in some sense any moral theory can be represented by a conse-

---

8.   Buchak (2013). Buchak's qualified departure from *Linear Discounting* is, I think, the most promising alternative to it. She argues that we have a distinct doxastic attitude—an attitude to risk—which is most felicitously captured by introducing a separate factor into the expected utility calculation. Although I disagree with Buchak's conclusion, I think this is the right way to approach this question.

quentialist counterpart.[9] I think that consequentialisers are probably right that any moral theory can be represented as giving a ranking on outcomes. I think they're wrong that we can represent all plausible views with a maximising decision rule— this is true, at least, if we aim to represent those theories not only for full information cases, but also for decision-making under risk (Lazar 2017a).

However, the detour into consequentialising may be unnecessary here. The simple presentation of decision theory that I gave above said that we should evaluate outcomes, and, for consistency, I stated *Ranking* and *Linear Probability Discount* in those terms. But outcomes are not truly fundamental to decision theory. What really matters is the level of normative support an act has, along some dimension, given that some state of the world is the case. It helps to understand this in terms of assigning utilities to the outcome of the act, given that state. But we could just as readily ask about the strength of reason in favour of that act, given that state. Deontologists should be perfectly comfortable saying whether one act is more or less supported by reason than another, holding the state constant, or else comparing act–state pairs. Indeed, if we can't say this, then I don't know how we can explain our choices between options.

Deontologists might still be worried about this framing, however. Doesn't it imply that I should be prepared to commit one *prima facie* wrongful act-type, if by doing so I can prevent others from performing a greater number of such act-types? This risk can be catered for in a number of ways. My preferred approach is to recognise that the reasons for an act can be agent-relative as well as agent-neutral.

Agent-relative reasons either apply, or have particular force, for specific agents. Agent-neutral reasons apply universally, and have the same force for everyone. So, if my son is drowning, then everyone has a strong agent-neutral reason to save him. But I also have a more powerful agent-relative reason to do so. As is now well established, by recognising agent-relative reasons, we can accommodate the core deontological idea that there are some constraints that we must not breach, even if our doing so would prevent more breaches of the same duty by others (McNaughton and Rawling 1995). We can also do justice to special duties grounded in our valuable relationships. I must attend, first and foremost, to my own duties, rather than to the goal of maximising duty-compliance overall. Once we take agent-relative reasons into account, breaching a duty to minimise duty breaches overall is not, in fact, your best option.

Agent-relative reasons allow us to account for *constraints*—the requirement not

9.   For an overview of the debate, see Portmore (2009).

to bring about what seems on the face of it to be the best outcome. But deontologists are also committed to options to act suboptimally. How can vanilla decision theory accommodate those?

It cannot. I doubt whether we can make sense of options to act suboptimally without recognising, as alluded to above, that there are distinct dimensions of normative strength. Vanilla decision theory presupposes that there is only one dimension of normative strength, which in turn explains why it is committed to *maximising* expected utility. If there is only one dimension of normative strength, and one option is stronger than another on that dimension, then there is simply no rational explanation why one would choose the lesser option.[10] By contrast, once we recognise that there are different dimensions of normative strength, this result is entirely predictable.

As with everything, there is some disagreement about precisely how to understand the different dimensions of normative strength.[11] My approach is quite simplistic. Sometimes a reason counts in favour of an act, without making it the case that not performing that act would be wrongful. That's the justifying dimension of normative strength. Sometimes a reason counts in favour of an act, in such a way that if you do not perform the act, you would be acting impermissibly. That's the requiring dimension of normative strength.

To illustrate this difference, consider a trio of trolley cases. In the first scenario, the trolley is headed towards ten people, whom I can save only by diverting it towards two people on another track. I am permitted to turn the trolley.[12]

In the second scenario, there are two tracks, with two levers. I can pull only one lever. On track A, trolley A is headed towards ten people, whom it will kill unless I pull lever A. On track B, trolley B is headed towards my son, whom it will kill unless I pull lever B. Pulling either lever diverts the relevant trolley down a side track where it stops harmlessly. I am permitted to turn trolley B.

In the third scenario, the trolley is again headed towards my son, and I can save him only by diverting it down a side track where it will kill two people. I am not permitted to turn the trolley.

---

10.   One might endorse 'satisficing', but one of the main objections to that approach is that it is straightforwardly irrational. There are other potent objections too: see Lazar (2017a).

11.   The distinction was introduced by Joshua Gert in the context of practical rationality, but also applies to moral reasons. I disagree with Gert's understanding of the distinction in various respects, but will not explore them here. See Gert (2007).

12.   This example adapts one that I introduced in Lazar (2013).

Let's assume that you share my judgements on these cases. If there were only one dimension of normative strength, then that would make us both incoherent—unless we argue, implausibly, that each case is simply very finely balanced, so you can go either way (we could fix that by changing the numbers—I think the verdicts would be robust over quite a lot of variation).[13] If saving ten *is* more important than killing two, and saving my son is more important than saving ten, then saving my son should also be more important than killing two.

Once we recognise that normative strength varies along two dimensions, however, this apparent incoherence goes away. On one dimension, saving ten is morally more important than not killing two, because saving ten can justify killing two. On another dimension, not killing two is morally more important than saving ten, because you can be required to let your child die, if that is necessary to avoid killing two people, whereas you cannot be required to let your child die in order to save ten.

Vanilla decision theory has no room for these subtleties. It recognises only one dimension of normative strength. The only remedy is to change the decision rule and, in particular, to reject its simple maximising framework. There are many different possibilities here, and adopting one rather than another may involve endorsing a particular account of what explains the requiring/justifying distinction. I don't pretend to have nailed that distinction, but as a proof of concept, here is an attempt that can capture a considerable range of the landscape.

> COST: *An act is subjectively permissible if and only if, and because, either (a) there is no all things considered expectedly better act or (b) every all things considered expectedly better act either (i) involves unreasonable marginal expected costs to the agent or (ii) is better only in virtue of expected benefits to the agent.*

The concept of an 'expectedly better act' maps directly on to what vanilla decision theory would describe as an expected-utility *ranking*. One act is expectedly better than another insofar as it realises more expected utility. We calculate an act's expected utility by considering all the possible agent-relative and agent-neutral reasons that would tell for and against that act, given the different possible ways the world might be, and then discounting them for the probability of their associated state coming about. The 'expectedly best' act is always permissible.

13.   You could appeal to vagueness to reach the same verdict, but that too is just a dodge.

This already gives us the machinery to accommodate most of the deontological credo described above. The constraint against violating fundamental rights is captured by that reason's agent-relative dimension. The same is true for our special duties. The role of intentions, beliefs, and the causal structure of harms and benefits can likewise be accounted for within our agent-relative and agent-neutral reasons.

Deontological constraints would be relatively easy to accommodate even within vanilla decision theory. But, as we have seen, its maximising rule elides the requiring/justifying distinction, so fails to account for options to act suboptimally. COST restores those options, and in so doing takes a particular approach to understanding justifying versus requiring. In COST, the moral betterness ordering captures the justifying dimension of normative strength. The expectedly best option is *always* justified. COST then captures the requiring dimension of normative strength by reflecting on the special authority we have over our own interests. In clause b(i) it states that, while an expectedly better option may be justified, it is not required unless the additional moral benefit is great enough to make the additional cost to the agent *reasonable*. And clause b(ii) states that if the better option is better only in virtue of how it serves the interests of the agent, then the agent cannot be required to pursue it.

This gives a coherent and explanatory account of the requiring/justifying distinction, allowing deontologists to adopt the central lessons of vanilla decision theory without abandoning options to act suboptimally. Some deontologists, however, might reject this attempt to explain the distinction by appealing to personal cost, seeing the latter as just one symptom of the requiring/justifying distinction, rather than the whole story. They might, then, favour a more schematic decision rule, along these lines:

> *RJ: An option is subjectively permissible if and only if, and because, either (a) there are no probability-weighted requiring reasons not to do it, or (b) the probability-weighted requiring reasons not to do it are outweighed by the combination of probability-weighted justifying and requiring reasons in its favour.*[14]

RJ is consistent with COST, but not exhausted by it. It could also license the view, for example, that in standard trolley cases one is merely permitted, not required,

---

14.   In formulating this principle, I was greatly helped by discussions with Tom Hurka about pro tanto duties and pro tanto permissions. And in an excellent paper, Daniel Muñoz has independently reached a very similar principle, for decision making under certainty: Muñoz (2018).

to turn the trolley (saving the five lives might justify killing the one, but not make it required). [15] It still draws on vanilla decision theory, insofar as it weights one's reasons in linear proportion to probability, and of course relies on the thesis that reasons can be weighed. It can still incorporate agent-relative and agent-neutral reasons, in just the same way COST can. Of course, RJ is more schematic than COST, and therefore less informative. But it does offer a proof of concept that one can extend orthodox decision theory to accommodate the key positions in the deontological credo, without doing any violence to the requiring/justifying distinction.

There remains at least one relatively common (though not uncontroversial) deontological precept that this criterion of subjective permissibility might not adequately accommodate. Plenty of deontologists think that it is wrong to trade off lives against headaches: there is no number of headaches that one could avert that would make it permissible to avert those headaches rather than save a single life. [16] Extending this kind of view to risky choices is extremely difficult.

The main problem can be articulated quite simply: deontological hostility to aggregation has to do with *how we weigh our reasons*. If the more important class of reason is at stake, the less important class of reason cannot even be counted. But the decision-theoretic approach has us total the expected reasons in favour of each act, and those reasons must account for all the relevant ways the world might be. Some of the ways the world might be include scenarios in which the higher-weighted reasons are not, in fact, actual. In those scenarios, the lower-weighted reasons can count. We then have to discount those reasons for the probability of that scenario arising. But once we have done so, we have an expected utility like any other, and it goes into the mix, to weigh against the expected utility associated with the higher-weighted consideration.

Suppose, for example, that you have to choose between treating someone who is at risk of dying and treating some number of people at risk of suffering a headache. If the person will certainly die if you do nothing, then you have no reason to save those who may suffer a headache. But he might not die. So there are possible states of the world where he won't die if you treat those with headaches. But if that's right, then we can weigh the expected utility of treating the headaches against the expected

---

15.   Thanks to Peter A. Graham for raising such cases with me. For further discussion of them, see our paper, Lazar and Graham (2019). For the contrary view about 'lesser evil options', see Frowe (2018).

16.   For the state of the art, see Tomlin (2017); Voorhoeve (2014).

utility of saving the life, even though we are not allowed to weigh actual headaches against actual lives.

There are ways to fix this problem, with an exotic theory of the weight of our reasons, or by articulating a complicated principle that accommodates hostility to aggregation. I develop those in detail elsewhere (Lazar 2018a). An alternative option of course is to reject the deontological hostility to aggregation.

6. WHERE NEXT?

I think the prospects for deontological decision theory are good. We can accommodate all of the central deontological tenets within a relatively simple decision rule—indeed, one that is only a little more complex than vanilla decision theory. Anti-aggregationism may be a challenge to accommodate, but not impossible, and anti-aggregationism itself faces serious questions (Tomlin 2017). And we have positive reason to endorse this decision rule, since *Ranking* and *Linear Discounting* seem to be very well supported.

Of course, RJ and COST are not the only alternatives. RJ is more general, COST is more partisan. We could propose other alternatives that account for the requiring/justifying distinction differently. Much of the work would go into figuring out how to enumerate and weigh the many different reasons that might apply to our actions. But that, of course, is simply the inescapable task of moral philosophy in general. More pressing, perhaps, are four additional kinds of question that deontological decision theory—and indeed any moral decision theory—must address.

First, which probabilities are relevant for subjective permissibility? Perhaps it depends on what we're trying to do. Quite plausibly, we can care about all of the following: the agent's actual probabilities; what her probabilities would be if her beliefs were consistent; the probabilities justified by the evidence she actually had; and the probabilities justified by the evidence she could have had, had she done more research. I think that we hold people to high epistemic standards, and that we do so on moral grounds: the relevant probabilities are those that you would have if you did the morally appropriate research. But that might involve some problematic circularity (how do we work out what amount of research counts as morally appropriate?) (Zimmerman 2008), and it might also contravene my earlier argument that practical matters should be addressed through our practical reasons, rather than through the probabilities.

Second, how should moral philosophers approach attitudes to risk? For example, suppose a villain presents me with the following choice. I can either flip a coin or roll a die. If I flip the coin and it lands heads, he will kill Alf and Betty; if tails, he kills nobody. If I roll a die and it comes up even, then he will kill Alf; if odd, he will kill Betty. If I refuse, he will kill us all. If I am risk-averse, then I should give additional weight to the worst-case scenario. This means I should roll the die, because its worst outcome is better than the worst outcome from tossing the coin. If I am risk-seeking, then I should give more weight to the best-case scenario, and so flip the coin, since it's the only option with the possibility of saving everyone. If I am risk-neutral, then I should plausibly be indifferent between coin and die.[17]

Recent work in decision theory has seen the resurgence of non-expected-utility theory, including ambitious attempts to build non-neutral attitudes to risk into decision theory from the ground up.[18] These philosophers tend to argue that it is sometimes permissible to adopt a non-neutral attitude to risk. That seems fine for rational decision theory. But if we are moral realists, then we must think that moral decision theory is more prescriptive. So, for any given choice, is there a morally appropriate attitude to risk that we should take?

One issue, which I have already addressed, is whether risk attitudes should be reflected in the utility, the probability-discount, or some third factor. But that is basically just a modelling choice. The more substantive question is just what attitude to risk one ought to take.

Scant work exists on this topic.[19] One natural view is that when one's action affects only oneself, one may adopt whatever reasonable attitude to risk one chooses, but when it affects others, one must instead act on their attitude to risk (Altham 1983). Lara Buchak argues that, if they have different risk attitudes, we should defer to the most risk-averse (within the reasonable range) (Buchak 2016b). I'm not sure about this. I think the risk attitude we adopt should also depend on what the stakes are—we should be more risk-averse when the worst-case scenario is really bad.

Third, just as people often have non-neutral attitudes to risk, so they also tend to prefer to decide on what they know, rather than act on ambiguous or impoverished information (Bradley 2016). We often face situations where one option has relatively

---

17.   This is actually not quite right; one could be risk-neutral but still care about the distribution of harms, so view the two options as different on those grounds.

18.   See especially Buchak (2013).

19.   Altham (1983); Buchak (2016b). It's also often easily confused with other distinct, but related topics—that's especially true in Spector (2016).

sharply defined probabilities, while another is much less precise; we tend to go for the more clear-cut alternative, even if the stakes of the other are higher.

For example, suppose you're a soldier deciding whether to go to war. You know that your friends are going, and that their lives will be at serious risk. You are a very talented soldier, so it is very likely that, if you go to war, you will at some point save some of your friends from suffering serious harm. But you also know that there's a chance the war is unjust, and that if it is unjust, then fighting is very seriously wrong, since every act of killing that you commit is the equivalent of murder (and can't be justified by the importance of saving your friends). You don't know enough, however, to assign any sharp probability to whether the war is just or unjust. It depends on information that neither you, nor anybody in the public, can find out. Many people would argue that you should give somewhat more weight to the outcomes about which you are more precisely confident (e.g. Betz 2016). You know that if you don't go, you will fail to save your friends from serious risks. That clear knowledge is worth something (Al-Najjar and Weinstein 2009; Voorhoeve et al. 2016).

The fourth topic is not, strictly speaking, one for moral decision theorists alone, but it does have special importance for us, and especially for deontological decision theorists. In many situations, if we assess our options in isolation from one another, we reach one verdict on their permissibility, but if we consider them as a set, we reach a different verdict (McClennen 1990). There are also interaction effects between choices within a series: sometimes whether it is permissible to perform an act depends either on what one did earlier, or on what one will do later. I have touched on this topic in a number of papers, arguing, for example, that sometimes moral 'sunk costs' should have a bearing on what it is permissible to do now (Lazar 2018C). I have also argued that whether we should assess individual options in isolation or as a campaign can depend on whether they are a necessary part of a series that is justified in the aggregate (Lazar and Lee-Stronach 2018). And I've considered cases in which one can seemingly be permitted to perform some suboptimal beneficent act when it is parcelled out into individual choices, but not when those choices are taken as a whole (Barry and Lazar 2018).

All of these cases suggest that moral philosophy needs a theory of dynamic choice—how to do the right thing over time, over a sequence of choices, not just one choice at a time. This is especially pressing for theories of moral decision making under risk, since choices with one risk profile when considered in isolation will have quite different risk profiles when taken as a sequence.

REFERENCES

Aboodi, R., A. Borer, and D. Enoch (2008) 'Deontology, Individualism, and Uncertainty', *Journal of Philosophy*, 105/5: 259-72.

Al-Najjar, N. I. and J. Weinstein (2009) 'The Ambiguity Aversion Literature: A Critical Assessment', *Economics and Philosophy*, 25/3: 249-84.

Altham, J. E. J. (1983) 'Ethics of Risk', *Proceedings of the Aristotelian Society*, 84: 15 - 29.

Barry, C. and S. Lazar (2018) 'Moral Options and Optimisation', Unpublished MS.

Betz, A. (2016), *The Ethics of War and Friendship: The Moral Significance of Fellowship of Arms*, (University of Illinois, Chicago).

Bolinger, R. J. (2017) 'Reasonable Mistakes and Regulative Norms: Racial Bias in Defensive Harm', *Journal of Political Philosophy*, 25/2: 196-217.

Bradley, R. (2016), 'Ellsberg's Paradox and the Value of Chances', *Economics and Philosophy*, 32/2: 231-48.

Briggs, R. (2014) 'Normative Theories of Rational Choice: Expected Utility', in *Stanford Encyclopaedia of Philosophy*, Edward Zalta (ed.).

Buchak, L. (2013) *Risk and Rationality*, Oxford: Oxford University Press.

———(2016a) 'Decision Theory', in *Oxford Handbook of the Philosophy of Probability*, Alan Hájek and Christopher Hitchcock (ed.), Oxford: Oxford University Press.

———(2016b), 'Taking Risks Behind the Veil of Ignorance', *Ethics*, 127/3: 610-44.

Colyvan, M., D. Cox, and K. Steele (2010) 'Modelling the Moral Dimension of Decisions', *Noûs*, 44/3: 503-29.

Dancy, J. (2004) *Ethics without Principles*, Oxford: Clarendon Press.

Ferzan, K. K. (2005) 'Justifying Self-Defense', *Law and Philosophy*, 24/6: 711-49.

Frowe, H. (2018) 'Lesser-Evil Justifications for Harming: Why We're Required to Turn the Trolley', *The Philosophical Quarterly*, 68/272: 460-480.

Gardiner, G. (2017) 'In Defence of Reasonable Doubt', *Journal of Applied Philosophy*, 34/2: 221-41.

Gert, J. (2007) 'Normative Strength and the Balance of Reasons', *The Philosophical Review*, 116/4: 533-62.

Haque, A. A. (2012) 'Killing in the Fog of War', *Southern California Law Review*, 86/1: 63-116.

Isaacs, Y. (2014) 'Duty and Knowledge', *Philosophical Perspectives,* 28/1: 95-110.

Kamm, F. M. (1985) 'Supererogation and Obligation', *Journal of Philosophy*, 82/3: 118-38.

Lazar, S. (2013) 'Associative Duties and the Ethics of Killing in War', *Journal of Practical Ethics*, 1/1: 3-48.

———(2015) 'Risky Killing and the Ethics of War', *Ethics*, 126/1: 91-117.

———(2017a) 'Deontological Decision Theory and Agent-Centred Options', *Ethics*, 127/3: 579-609.

———(2017b) 'Anton's Game: Deontological Decision Theory for an Iterated Decision Problem', *Utilitas*, 29/1: 88-109.

———(2018a) 'Limited Aggregation and Risk', *Philosophy & Public Affairs*, 46/2: 117-59.

———(2018b) 'In Dubious Battle: Uncertainty and the Ethics of Killing', Philosophical Studies, 175/4: 859-83.

———(2018c) 'Moral Sunk Costs', *The Philosophical Quarterly*, 68/273: 841–61.

Lazar, S. and C. Lee-Stronach (2018) 'Axiological Absolutism and Risk', *Noûs*, 53/1: 97-113.

Lazar, S. and P. A. Graham (2019) 'Deontological Decision Theory and Lesser-Evil Options', *Synthese*.

MacAskill, W. (2013) 'The Infectiousness of Nihilism', *Ethics*, 123/3: 508-20.

McClennen, E. F. (1990) *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge: Cambridge University Press.

McNaughton, D. and P. Rawling (1995) 'Value and Agent-Relative Reasons', *Utilitas*, 7/1: 31-47.

Muñoz, D. (2018) 'Better to Do Wrong', Unpublished MS.

Nair, S. (2014) 'A Fault Line in Ethical Theory', *Philosophical Perspectives*, 28/1: 173-200.

Oddie, G. and P. Milne (1991) 'Act and Value: Expectation and the Representability of Moral Theories', *Theoria*, 57/1-2: 42-76.

Olsen, K. (2018) 'Subjective Rightness and Minimizing Expected Objective Wrongness', *Pacific Philosophical Quarterly*, 99/3: 417-41.

Portmore, D. W. (2009) 'Consequentializing', *Philosophy Compass*, 4/2: 329-47.

Quong, J. (2015) 'Rights against Harm', *Aristotelian Society Supplementary Volume*, 89/1: 249-66.

Smith, H. (2018) *Making Morality Work*, Oxford: Oxford University Press.

Smith, N. J. J. (2014) 'Is Evaluative Compositionality a Requirement of Rationality?', *Mind*, 123/490: 457-502.

Spector, H. (2016) 'Decisional Nonconsequentialism and the Risk Sensitivity of Obligation', *Social Philosophy and Policy*, 32/2: 91-128.

Temkin, L. S. (2012) *Rethinking the Good : Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.

Tenenbaum, S. (2017) 'Action, Deontology, and Risk: Against the Multiplicative Model', *Ethics*, 127/3: 674-707.

Thomson, J. J. (1986) *Rights, Restitution, and Risk: Essays in Moral Theory*, Cambridge, Mass.: Harvard University Press.

Tomlin, P. (2017) 'On Limited Aggregation', *Philosophy & Public Affairs*, 45/3: 232-60.

Voorhoeve, A. (2014) 'How Should We Aggregate Competing Claims?', *Ethics*, 125/1: 64-87.

Voorhoeve, A., et al. (2016), 'Ambiguity Attitudes, Framing, and Consistency', *Theory and Decision*, 81/3: 313-37.

Zimmerman, M. J. (2008) *Living with Uncertainty: The Moral Significance of Ignorance*, Cambridge: Cambridge University Press.

Zollman, K. J. S., A. Bjorndahl, and A. J. London (2017) 'Kantian Decision Making under Uncertainty: Dignity, Price, and Consistency', *Philosophers' Imprint*, 17/7: 1-22.