Fairness in Philosophy
CS 256, Algorithmic Fairness, Stanford, Omer Reingold, 2023

Our course is about the TOC (theory of computing) foundation of algorithmic fairness, but the notion of fairness is much older and thus we will ground our discussion in fairness as other disciplines view it, with special attention to fairness in political philosophy. The following summaries are at a very high level and can be viewed as a philosophy-for-non-philosophers introduction. The term summaries were written by Olga Lenczewska and Cesar Manuel Valenzuela Marquez at Stanford University as part of the Simons Collaboration on the Theory of Algorithmic Fairness. The terms are meant to be readable in isolation but are obviously related to each other.
<table>
<thead>
<tr>
<th>Models of Affirmative Action</th>
<th>..........................................................</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compensatory Model</td>
<td>........................................................................</td>
<td>18</td>
</tr>
<tr>
<td>Diversity Model</td>
<td>........................................................................</td>
<td>19</td>
</tr>
<tr>
<td>Discrimination-Blocking Model</td>
<td>.......................................................................</td>
<td>19</td>
</tr>
<tr>
<td>Integrative Model</td>
<td>........................................................................</td>
<td>19</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>20</td>
</tr>
<tr>
<td>Power Dynamics</td>
<td>........................................................................</td>
<td>21</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>22</td>
</tr>
<tr>
<td>Deontological Ethics</td>
<td>........................................................................</td>
<td>23</td>
</tr>
<tr>
<td>Varieties of Deontological Theories</td>
<td>..................................................................</td>
<td>23</td>
</tr>
<tr>
<td>Strengths of Deontology</td>
<td>........................................................................</td>
<td>24</td>
</tr>
<tr>
<td>Relation to Fairness and Challenges</td>
<td>..................................................................</td>
<td>25</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>25</td>
</tr>
<tr>
<td>Consequentialist Ethics</td>
<td>........................................................................</td>
<td>26</td>
</tr>
<tr>
<td>Moral Assessment of Outcomes</td>
<td>......................................................................</td>
<td>26</td>
</tr>
<tr>
<td>Total Consequentialism</td>
<td>........................................................................</td>
<td>26</td>
</tr>
<tr>
<td>Act vs. Rule Consequentialism</td>
<td>....................................................................</td>
<td>28</td>
</tr>
<tr>
<td>Challenges to Consequentialism</td>
<td>...................................................................</td>
<td>28</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>29</td>
</tr>
<tr>
<td>Responsibility</td>
<td>........................................................................</td>
<td>30</td>
</tr>
<tr>
<td>Collective Responsibility</td>
<td>........................................................................</td>
<td>30</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>31</td>
</tr>
<tr>
<td>Egoism</td>
<td>........................................................................</td>
<td>33</td>
</tr>
<tr>
<td>Psychological Egoism</td>
<td>........................................................................</td>
<td>33</td>
</tr>
<tr>
<td>Ethical Egoism</td>
<td>........................................................................</td>
<td>34</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>35</td>
</tr>
<tr>
<td>Freedom and Liberty</td>
<td>........................................................................</td>
<td>36</td>
</tr>
<tr>
<td>Utilitarian Framework of Liberty</td>
<td>..................................................................</td>
<td>36</td>
</tr>
<tr>
<td>Liberal Egalitarianism</td>
<td>........................................................................</td>
<td>37</td>
</tr>
<tr>
<td>Libertarian Framework</td>
<td>........................................................................</td>
<td>37</td>
</tr>
<tr>
<td>Negative vs. Positive Liberty</td>
<td>.....................................................................</td>
<td>37</td>
</tr>
<tr>
<td>Further Reading</td>
<td>........................................................................</td>
<td>38</td>
</tr>
<tr>
<td>Cooperation</td>
<td>........................................................................</td>
<td>39</td>
</tr>
</tbody>
</table>
Fairness

Fairness refers to the impartiality in the rules and treatment that individuals give to one another in interactions where a certain benefit or burden is to be distributed. A common example of fairness is to divide a cake in equal pieces when the cake has to be shared by different people. It would be unfair for the person who cuts the cake to take a greater piece, or give better portions to their friends. The notion has played a prominent role in political philosophy through the work of John Rawls (1999; 2001). Rawls defined justice (a virtue of the most fundamental and pervasive institutions in society) as fairness. This means that the terms of social cooperation ought to meet two conditions (2001, 6):

First, the terms are reasonable enough to be accepted by everyone, and sometimes citizens are required to accept them if everyone else does. In other words, the principles that organize society should be justifiable to everyone who must honor them. Such justification is impossible when the rules of social cooperation are biased in favor of a group based on their natural talents, socioeconomic position, generation, conception of the good, race, or gender. All those criteria are taken to be arbitrary in the design of the fundamental principles of society and, thus, social arrangements that are guided by such criteria are unfair.

Second, for the terms of cooperation to be fair, they must also articulate an idea of reciprocity, that is, the benefits of cooperation are to be enjoyed by all those who do their part as stipulated by the rules. This condition precludes situations of exploitation and domination. If a group in society is allowed to extract disproportionate benefits from other groups by taking advantage of an arbitrary form of power (e.g., men's reliance on women's domestic labor in order to pursue their own life projects), then the terms of cooperation are unfair.

Fairness and Equality

Although fairness does not mandate strict equality (e.g., wage inequalities can be fair if the labor market is properly settled), it does require the rules of the game, as it were, to be designed such that participants stand in a relation of equality in some sense. A game where a group of participants has access to relevant information that others lack or is entitled to change the rules to their convenience, is unfair. In a democracy, this means
that citizens must have equal say in the organization of their society and be able to object to outcomes they find unjustifiable.

Further Reading

Justice

Justice is one of the key concepts of political philosophy—a discipline that reflects on the questions of how to best regulate social life through public institutions and legal norms. To say that an action is demanded by justice is to say that it is obligatory (as opposed to supererogatory) and that it is enforceable through a system of institutional rewards and punishments. This enforceability is what distinguishes justice from morality, the rules of which do not have to be enforced through law. While rules of justice are limited to regulating relationships between citizens or residents (between members of a political community), rules of morality are broader in scope and include the regulation of family relations, friendships, and even one’s self-improvement. Traditionally, then, justice has applied to the “public sphere” of life, while morality to the “private sphere”—although this distinction has been criticized by some contemporary frameworks, in particular by feminist theorists who believe that certain family structures may be themselves unjust. While to act justly typically means to abide by existing just legal rules, such rules can themselves be unjust; in this case, justice demands not that we abide by existing legal norms, but that we abolish or reform them.

Justice vs. Fortune

Another useful contrast to draw is that between justice (or injustice) and fortune (or misfortune): while the former describes states of affairs that an agent has contributed to bringing about (either by intentional action or omission), the latter refers to states of affairs that came about without human involvement—for example, natural disasters. While various theories of justice have been proposed by philosophers in the past millennia (going back as far as to Plato), and while there is no consensus about the content of the rules of justice, philosophers typically agree that the concept of justice only applies to an environment of moderate scarcity characterized by people having conflicting claims on existing resources (such as wood or water), which consequently need to be distributed according to some principle. By contrast, justice does not apply in an environment abundant in resources, in which everybody can take what they want and there is plenty left for others, or in an environment of extreme scarcity, in which there are not enough resources even to fulfill everyone’s basic needs.

Distributive and Retributive Justice

Existing theories of justice may be broadly divided into “distributive” and “retributive.” Distributive justice regulates existing relationships between all the members of the community—relationships assessed through a “currency” of justice (such as equality of opportunity) specifying in which respect justice requires that people are equally well off.
Retributive justice, by contrast, regulates relationships between people by looking at their past; for example, it can regulate a relationship between an offender and a victim, requiring that the offender remediates the wrong he committed in the past.

Further Reading

- [https://plato.stanford.edu/entries/justice/](https://plato.stanford.edu/entries/justice/)
- [https://justiceharvard.org/](https://justiceharvard.org/)
Equality

Equality is the chief political principle of democracy. A form of government is democratic only if citizens' interests (provided that they are reasonable) enjoy equal respect, consideration, and representation. It is widely accepted that universal franchise and equal treatment before the law are necessary to that end. However, contemporary theories of equality (generally grouped under the label "egalitarianism") have also emphasized the insufficiency of formal conditions to achieve political equality in any significant way. High levels of economic inequality, for instance, may lead to the overrepresentation of the wealthiest groups of society in decision-making bodies. To address these issues, the scope of egalitarianism has broadened in order to examine a variety of institutions in social life (the market, the workplace, the family, the school, etc.) and determine which inequalities are unjust or simply incompatible with democracy.

Theories of Equality

Since equality is a three-element relation between at least two entities and the feature with respect to which they are being compared, a theory of equality needs to answer two basic questions: "equality among whom?" and "equality of what?". While moral equality among all human beings is presupposed, theorists have focused on the equality that holds among citizens of the same country and differ more drastically on their answer to the second question. Nonetheless, recent debates about immigration, global justice, intergenerational justice, and animal rights, have led to more fine-grained accounts of who ought to be equals.

Several answers are available to the "equality of what?" question.

Equality of Outcomes

The most radical notions of equality entail some level of equality of outcomes: over the course of their lives, everyone gets the same amount of material goods or experience the same levels of welfare. The equalization of material goods is often associated with some brands of communism and socialism. Certain forms of market socialism, for example, propose equality of post-tax income. Authors in these traditions tend to emphasize how differences in income and wealth translate into differences in political and social power, and claim that any form of substantive equality is simply incompatible
with economic inequalities. Equality of welfare, on the other hand, is closer to utilitarianism and seeks to distribute resources differentially in order to guarantee equal levels of satisfaction and dissatisfaction among people. Both forms of egalitarianism are usually rejected in liberal accounts of equality. Equality of material goods would impose a unique conception of what is valuable, whereas the liberal state is supposed to remain impartial vis-à-vis the different and reasonable conceptions of the good that emerge in a pluralistic society. Likewise, equality of welfare might lead to the mobilization of resources for the satisfaction of unreasonable preferences (e.g. a taste for highly expensive wine), which are seen as illegitimate by the liberal state.

Equality of Opportunity

A more consensual type of equality is equality of opportunity. This ideal requires that positions of power are open to everyone (i.e. there cannot be any formal discrimination) and - what is far more demanding - that people with equal talents and equal willingness to use those talents have equal chances to attain those positions (the second condition is defended by John Rawls, 1971/1999). Equality of opportunity has the advantage of being deeply embedded in the culture of Western capitalist societies. The promise of the so-called "American dream," for instance, was precisely that hard work would amount to better outcomes, regardless of who you were or where you came from. Nevertheless, there is a debate whether people need equal opportunities only in the early stages of their lives (e.g. equal access to K-12 public education) or across their lifespan (e.g. equal access to healthcare during adult life). This is the difference between starting-gate theories and lifespan theories of equality.

Moreover, some inequalities might still be objectionable even when they arise among people who have equal opportunities. Under initially equal conditions and in a perfectly meritocratic society, a group could gain so much economic or social power that it is able to subjugate others, and such power would be objectionable.

Equality of Resources

Equality of opportunities is linked to equality of resources to live a good life. The list of resources needs to be broad enough to ensure that citizens with many different conceptions of the good can pursue their life plans. At the same time, the distribution should not depend on the particular levels of welfare that each individual would extract from her share of resources. One approach, formulated by Rawls, is to conceive
resources as primary social goods, that is, goods that every rational person would want and that are necessary to be a cooperating member of society. Such goods are rights, liberties, opportunities, income and wealth, and the social bases of self-respect. For Rawls, inequalities in such goods are acceptable only if they are to the greatest benefit of the least advantaged.

What if inequalities arise from the way people use their resources? A frequent concern among egalitarians is how to deal with the question of responsibility. Some have argued that inequalities are justified insofar as they are the predictable results of people's deliberate choices (e.g. the deprivation experienced by a surfer in Malibu who chooses not to work). This means that we should restore equality only when the unequal outcomes were unavoidable, unforeseeable or unlikely (e.g. being hit by a drunk driver).

Relational Egalitarianism

A different focus is proposed by relational egalitarianism (pioneered by Elizabeth Anderson 1999). Relational egalitarians reject the idea that equality is only or even primarily about distributing goods, or that we should be worried about correcting the "injustices" of luck. They emphasize that equality is a relation between citizens that enjoy equal status (despite differences in income and wealth), treat each other with respect, are free from domination and oppression, and have access to the resources necessary to be functioning members of society. Social hierarchies of any kind are thereby rejected insofar as they impede citizens from relating as equals.

Value of Equality

A debate remains whether equality has intrinsic or instrumental value. The question is often triggered by the "leveling-down objection": if equality were intrinsically desirable, then it would be acceptable to solve inequalities by making the better off person worse off until she is in the same position as the person who was initially worse off, all things equal. This seems inefficient and wrong. Motivated by this concern, some have proposed other principles of distributive justice, like sufficiency or priority. Sufficientarianism argues that we should not worry that everyone has the same but, instead, that everyone has enough to remain above a threshold of decent quality of life. Prioritarianism, in turn, holds that the allocation of resources should be done by giving priority to the gains we can expect in the well-being of the worse off.
Oppression

Oppression is a fundamental kind of injustice where individuals or groups are subjected to a source of power that systematically puts them in a position of unfair and deep disadvantage. Victims of oppression face difficulties to develop their abilities, enjoy the benefits of social cooperation, pursue their life plans, express themselves, or be recognized as equals in their societies. In that sense, oppression affects different aspects of people’s lives and compounds a variety of injustices. This pervasiveness is what makes the term so powerful in public discourse. For instance, if we say that women are oppressed by men, we are probably emphasizing more than a slight disadvantage or the failure of a few legal procedures: we are pointing, on the one hand, to the wide variety of obstacles that constrain women’s development and autonomy in crucial aspects of their lives, and on the other hand, to the many ways in which men benefit from women’s vulnerability and seek to maintain it.

The paradigmatic case of oppression is tyranny: the ruler has arbitrary power over people and, with total disregard for their wellbeing, subjects them to the constant threat of violence. In this case, oppression is essentially a political problem. However, as Ann Cudd (2006) notes, this political conception has yielded since the nineteenth century to a more social and comprehensive one, where spheres like the economy, the family, or culture may also show instances of oppression.

Structural Oppression

Expanding the notion of oppression has led to the distinction between “agent oppression” and “structural oppression,” studied by Sally Haslanger. Agent oppression is inflicted by an identifiable person or group of persons, the oppressor, upon another, the oppressed. Here the oppression stems from the harshness of the harm and the fact that it’s an abuse of power (e.g. a police officer who deliberately harasses African American citizens in a local community). Structural oppression, instead, is imposed through social structures, that is, through the set of practices, institutions, and norms that distribute resources and constraints. In this case, Haslanger claims, the problem is the misallocation of power entailed by the institution or practice (e.g. the set of legal institutions, like the police, courts, and prisons, that widen and perpetuate racial
inequalities by making it far more likely for African Americans to be incarcerated than for European Americans).

**Defining Structural Oppression**

Although both types of oppression may be intentional or not (an agent could harm another either deliberately or unintendedly; a structure could be designed to disadvantage a group or may just have that unforeseen effect), both Haslanger and Iris Marion Young have claimed that structural oppression may obtain even when no moral agent is blameworthy or responsible for the emergence of the oppressive structure. For Young, oppression only entails the existence of a disadvantaged group and a group that is privileged with respect to the former, but the latter may or may not count as the correlating oppressor. These accounts often face two challenges: 1) explain how structural oppression could be a form of moral wrong even when there is no wrongdoer; and 2) offer an account of how responsibility to end oppression should be distributed among blameless members of a society. The debate remains open on both questions.

A key idea in accounts of structural oppression is that individuals are oppressed by virtue of being members of a certain social group. According to Young, social groups are collectives of persons who are differentiated from other groups by cultural norms and practices. Although membership is not voluntary, it shapes people’s identity insofar as it has a considerable impact in their lives. Now, in Haslanger’s terms, being a member of an oppressed group means that membership to that group is non-accidentally correlated with some form of disadvantage. In other words, the disadvantage of a set of individuals is explained by their membership to a specific social group. This does not imply that all members of the group are equally oppressed, or that they all experience exactly the same disadvantages, or that they are the only ones in society experiencing certain disadvantages. The degree of oppression depends on the other social positions each individual occupies. A wealthy woman, for instance, may have greater risk of suffering sexual violence than men but a lower risk than a poor woman; she may also earn more money than some men. None of this contradicts the idea that she is oppressed as a woman. Further, for the relevant correlation to obtain, policies need not explicitly target the oppressed group. For example, a policy may disadvantage Native Americans, not because it is directed at them, but because it disadvantages poor people, and due to prior injustice, there is a correlation between being Native American and being poor.
Forms of Structural Oppression

Young considered that structural oppression can take five forms, which can all exist independently but more often than not emerge together: *exploitation* (transfer of results of labor of one social group to another in order to augment the power, status, and wealth of the former); *marginalization* (a social group is excluded from useful participation in social life and subjected to material deprivation and/or extermination); *powerlessness* (members of the oppressed social group, usually non-professionals, lack avenues for recognition and respectability, have little or no work autonomy or technical expertise, and are impeded from exercising their judgment or creativity); *cultural imperialism* (one’s group perspective is rendered invisible, stereotyped, or marked as deviant or inferior); and *violence* (members of the oppressed group know that they face the systematic threat of physical harm).
Discrimination

Discrimination refers to a form of unequal treatment that unfairly disadvantages some persons with respect to others on the basis of their membership to a social group, that is, on the basis of their sex, gender, sexuality, race, origin, religious or political beliefs, physical appearance, economic status, or other similar grounds deemed arbitrary in the enforcement of people’s rights. For instance, Jane is discriminated against as a woman in her job application if a male candidate who is less qualified than her is hired and she is rejected because he is a man and she is a woman. This definition is normatively charged: here, discrimination is always a moral wrong. In a broader, neutral sense, we discriminate against people every time we give them differential treatment, e.g. when people who do not have a medical degree are impeded from applying to a position as a doctor. However, philosophers most commonly use the term in its pejorative sense. They also tend to agree that not only persons but also rules, practices, and institutions may discriminate (see Racism; Oppression), and that victims of discrimination are most commonly persons and social groups (e.g. African Americans, women, LGBTQIA, Muslims, etc.).

Discrimination versus Differential Treatment

Not every form of differential treatment that entails an unfair disadvantage is properly described as discrimination. For instance, if Marie is the mayor of her city and only deals construction contracts with members of her own family, she is treating contractors unequally and unfairly disadvantaging some of them. Yet her action would be better described as “nepotism,” not discrimination. This has led some philosophers to emphasize the “social salience” of the group that is being discriminated against: a group is socially salient when the features of its members are more or less publicly recognized as such and perception of those features has a pervasive impact in different aspects of social life (Lippert-Rasmussen 2014). “Being Hispanic” or “being transsexual” are socially salient categories, whereas “not being a member of Marie’s family” or “being a fan of horror movies” are not.

Likewise, unequal treatment may be morally justified and promote fairness. Defenders of affirmative action, for example, hold that unequal treatment that advantages members from oppressed groups compensates pervasive inequalities in opportunities (see Affirmative Action; Equality). Schools with a majority of African Americans students,
for instance, tend to be underfunded, which leaves these students in disadvantage with respect to White Americans in their college applications. In such cases, it may be justified for the applicants’ race to be considered when making admission decisions.

**Direct Discrimination**

An important distinction is made in philosophy and law between *direct* and *indirect* discrimination to account for the fact that discrimination does not depend on the psychological states (intentions, indifference, biases) of those who discriminate (Altman 2020; Lippert-Rasmussen 2014). *Direct* discrimination refers to cases where agents either design rules with the aim to disadvantage certain groups, or their practices reflect unconscious biases, indifference, or other reprehensible mental states towards the disadvantage of the affected groups. Jim Crow laws are a clear example of the first type of case: not only did they openly disadvantage African American in their access to public goods and spaces, but even when the letter of the law did not target Blacks explicitly (e.g. voting restrictions), their disadvantageous effect on this group was deliberate. An example of the second type of case, where other objectionable mental states like biases are at work, would be a department store that requires workers to wear a specific kind of hairstyle because it “attracts customers” but, in fact, is likely to exclude certain ethnic groups.

**Indirect Discrimination**

However, discrimination may also be *indirect*. This happens when a policy or law is facially neutral vis-à-vis social groups but, in practice, has disadvantageous (yet unintended) effects for some groups and such effects are attributable to the mental states of the agents implementing the policy. The classic example is the 1971 Supreme Court case *[Griggs v. Duke Power]*. In order to qualify for a promotion, a North Carolina company required its employees to take two tests, which Blacks were far less likely to pass than Whites. The court ruled that these requirements did not measure job-related skills and, given the state’s history of discrimination against Blacks, the policy was discriminatory. For there to be a moral harm in this case, one does not need to prove intent to discriminate against Blacks. This does not mean that claims of indirect discrimination are necessarily easier to prove; they are often contested in cases where other explanations for the disadvantage seem plausible.
Further Reading

Sexism

Broadly speaking, sexism is a form of discrimination or oppression on the basis of sex. In this context, to be discriminated means to face some form of differential (lesser) treatment, whether the discriminating agent intends this or not; and to be oppressed means to suffer some inhibition of the ability to develop and make use of one’s capacities as well as to express one’s thoughts and feelings.

Within philosophy, sexism is one of the key concepts in feminist philosophy and is frequently addressed in contemporary moral and political philosophy as well. Feminist philosophers aim to diagnose, describe, and provide solutions to sexist attitudes and behaviors, both those of individuals and those displayed by public institutions. Sexism is frequently discussed in the context of the notion of justice, as it is the primary phenomenon designed by “gender injustice” or “gender inequality”. This is because most philosophers believe that a discriminatory or oppressive treatment of women is fundamentally unjust (and it is also, by definition, unequal to the treatment of men).

While history provides numerous examples of explicit and intentional sexist attitudes on the part of men, the kind of sexism that socio-political philosophy and feminist philosophy are interested in nowadays is primarily implicit or unintentional. Few individuals in Western societies believe today that women are, say, inherently less intelligent or that they have an obligation to become full-time homemakers regardless of their personal preferences; yet those individuals might sometimes behave in sexist ways nonetheless. Similarly, while lawmakers rarely intend to enact laws and policies that discriminate against women, they can nonetheless possess some discriminatory features that limit women’s options in conjunction with other social structures (a phenomenon known as “structural sexism” or “structural injustice” more broadly). The implicit and unintentional nature of most instances of contemporary sexism makes it more difficult to notice it and, consequently, to address it.

The Private vs. Public Sphere

For the past 50 years, philosophical writings diagnosing and addressing sexism have focused on two broad domains of life where sexism occurs: first, at home or within the so-called “private sphere” of life; second, in the workplace and other public spaces or within the so-called “public sphere”. Inside the home, women frequently assume the
majority of care and housekeeping responsibilities, often to the detriment of their own employment prospects and financial security. This phenomenon by itself is not necessarily an instance of sexism. But relevant examples of sexism include societal attitudes and norms exhibited by a male partner, parents, or teachers, which pressure women to assume the majority of care- and housework, or to prioritize their male partner’s career over their own, or to internalize standards of behavior and bodily appearance to which men are not held. Within public contexts such as politics or employment, women may once again be held to different standards than men, their actions interpreted differentially because of their gender. In the workplace, in particular, the most pressing issues include the gender pay gap and discriminatory policies that disadvantage women who have, or want to have, children (phenomena often known under the umbrella term “pregnancy penalty” or “motherhood penalty”).

Because individual attitudes are often unconscious and unintended, feminist and political philosophers believe that sexism should be addressed primarily through large-scale, gradual changes in social norms, education patterns, and policies, instead of by individuals examining and changing their beliefs. Moreover, the large-scale challenge of addressing sexism is further complicated by ways in which sexism interacts with specific cultural contexts and ways in which it intersects with other forms of discrimination and oppression, such as racism or ableism.

Further Reading

https://plato.stanford.edu/entries/feminist-philosophy/

https://plato.stanford.edu/entries/feminist-power/

Racism

Racism is, primarily, a form of discriminatory treatment that unfairly disadvantages members of a racialized group in virtue of their membership to that group. Individuals in the targeted group are thereby relegated to a position of political, economic, social, and cultural subordination. Such unequal treatment may or may not be accompanied with an explicit conception of members of the targeted group as being intrinsically inferior, deviant, or simply less deserving of the benefits of social cooperation. More broadly, racism may just refer to the representation (through beliefs, discourse, symbols, etc.) of a racialized group as inferior, regardless of whether and how that representation leads to discriminatory treatment and disadvantageous effects.

Notions of Racism

Originally, the term “racism” was meant to apply to systems of beliefs that purported a racial hierarchy, e.g. the Nazi belief system (Bloom 2002), but this ideological use, both in philosophy and public discourse, has proven to be limited. Many different things and phenomena may be called “racist” now: not only beliefs and ideologies, but also attitudes, utterances, individuals, groups, societies, laws, institutions, symbols, etc. This expansion of the term is doubly motivated. On the one hand, individuals may hold racist beliefs and attitudes, and institutions and laws may produce racial inequalities, even if they are not articulated within a racist ideology. On the other hand, racial disparities have persisted and, in some respects, even worsened despite growing and widespread moral condemnation of racist ideologies.

Institutional Racism

The latter fact has led to the development in philosophy and social science of the notions of “institutional racism” and “structural racism.” These are supposed to capture the disadvantaging effects that institutions and social structures, respectively, may have on racialized groups, independently of whether they were designed to have such effects, or whether their participants intend those results or are consciously or overtly racist. In that sense, according to Tommie Shelby (2016), “institutional racism” refers to two ways in which an institution may be racist: 1) either the institution’s policies are formulated in race-neutral terms and yet have a disproportionate impact on a disadvantaged racialized group; or 2) the goals, rules, or application of procedures in the institution are,
respectively, justified through a racist ideology, racially biased, or discriminatory, or they are in the hands of individuals who are consciously or unconsciously prejudiced against a racialized group.

**Structural Racism**

“Structural racism” is a broader category and refers to the compounded disadvantageous effects that social structures, that is, sets of interconnected practices, institutions, and rules, have on racialized groups. As a form of structural oppression, structural racism is a phenomenon in which social processes systematically put members of a racialized group under conditions of unfair disadvantage, deprivation, or domination (see entry on Oppression). In such cases, the locus of racism is in the structures of society that misallocate power by privileging one racialized group and disadvantaging another. Although individuals retain some level of agency within social structures, they face both material and symbolic constraints to different extents (e.g. lesser levels of wealth, or lesser access to positions of power). By appealing to the notion of structural racism, philosophers aim to explain the fact that, despite there not being any genetic, psychological, or cultural differences among racialized groups that could account for the situations of deprivation experienced by some of them, such deprivation occurs and persists systematically throughout society.

Philosophers remain in disagreement over the necessary conditions to call something or someone “racist,” given the moral charge of the accusation. Some insist that even institutional or structural racism must be underpinned by racist *ideologies* that spread false beliefs about race, namely about the inferiority of some groups, however subtle, unconscious, or inconsistent those ideologies may be. This, of course, does not exempt individuals from the responsibility of endorsing or reproducing those ideologies. Others emphasize the presence of adverse emotions and attitudes, or the lack of respect, towards a racialized group as a necessary condition for something or someone to be racist. Most authors, though, would reject the idea that deliberate intent or design is necessary for a person, an action, or a thing to count as racist.
Affirmative Action

Affirmative action refers to a policy aimed at increasing participation of historically disadvantaged groups (namely, women, racial and ethnic minorities) in key institutions like government, universities, and firms. Institutions may engage in affirmative action in two ways: either by simply encouraging participation or by preferential selection, that is, taking membership to a disadvantaged social group as a criterion to select employees, students, contractors, etc.

The preferential form of affirmative action is highly polemical. Critics often claim that affirmative action is a form of reverse discrimination, that is, the policy unfairly disadvantages members of some groups (e.g. Whites, men) in virtue of their membership to that group, which should always count as a moral harm. In that sense, an uncompromised commitment to equality, civil rights, and the rule of law, the argument goes, would exclude any consideration of race, ethnicity or gender in the allocation of positions and offices. Critics also emphasize how affirmative action violates meritocratic principles by including and even prioritizing criteria of selection other than skills or performance. For some, this also fosters the stigmatization of disadvantaged groups as “undeserving.”

Models of Affirmative Action

The challenge for defenders of affirmative action is to show that the policy overall represents a fair benefit for members of disadvantaged groups while avoiding novel forms of racial discrimination. Philosophers, legal scholars, and courts in the United States have offered at least four different models to justify (race-based) affirmative action practice, summarized by Elizabeth Anderson (2010):

Compensatory Model

On the compensatory model, affirmative action is a way for society to compensate disadvantaged groups for the harm perpetrated against them in the past. For instance, given the pervasive effects of racial discrimination during Jim Crow, it is fair to assume that most African Americans were affected by that harm (without examining which individuals were actually affected, which would be impossible) and that the costs of compensating them ought to fall on society as a whole, even if some “innocent” people...
(e.g. young whites who are not selected in colleges engaging in affirmative action) are affected.

Diversity Model

According to the diversity model, affirmative action increases the diversity of ideas within an institution. Insofar as epistemic diversity is a legitimate end and the presence of different racial groups may foster that end, institutions are justified in considering race in their selection processes. Here racial diversity is only one relevant form of diversity among many others. Moreover, since only the best qualified members of disadvantaged groups would actually further epistemic diversity, the diversity model remains in tune with meritocratic principles.

Discrimination-Blocking Model

In contrast, the discrimination-blocking model conceives affirmative action as a tool to counteract the effects of current forms of discrimination that are no longer the product of blatantly discriminatory law but of persistent and widespread stigmas and biases. Affirmative action then adds pressure on institutions to eliminate discrimination, since they would otherwise tend to give free rein to surreptitious forms of discrimination.

Integrative Model

Finally, the integrative model argues that affirmative action reduces ongoing (de facto) segregation and stigmatization by ensuring that disadvantaged groups are properly represented in key institutions. Integration is a requirement of democracy and affirmative action is a means to achieve it: integrated institutions enable communication and cooperation between social groups and allow them to build a collective identity. Anderson herself defends this model over the others (148-154).

Each of these models remains controversial. Some issues surrounding affirmative action, though, are empirical, not just theoretical (see Fullinwider 2018). They mostly pertain to the positive and negative effects that affirmative action has on minority groups: does it guarantee greater participation, all things considered? Does it favor their academic or professional achievement? In that sense, defenders of affirmative action, in the face of increasing resistance, must draw on both moral and empirical arguments. But they also must acknowledge that a part of that resistance will not be undermined by rational
dialogue and, instead, requires them to address unconscious biases and recalcitrant prejudices.

Further Reading

Power Dynamics

Within political philosophy, the notions of power dynamics and bargaining power are relevant to normative or evaluative—as opposed to descriptive—theorizing about relations between people. In other words, the goal is not to observe and describe existing power relations within, say, the workplace—a task of behavioral and social sciences—but to develop an account of what power dynamics ought to be in order to be just and fair, and to critique existing social phenomena through such an evaluative framework. Broadly speaking, a just and fair way of managing power dynamics between individuals includes ensuring that each person starts on an equal footing (for example, everybody’s basic needs are acknowledged and everybody’s opinions are voiced) and that each person is presented with a variety of options and is free to choose how to proceed, including backing out of the situation or transaction in question if she so wishes.

The notions of power dynamics and bargaining power are relevant to all kinds of interpersonal interactions, including those occurring within the household, at the workplace, or in an educational setting. Power dynamics within the household have been a key focus of feminist philosophy and theory of the last 50 years. According to this school of thought, composed both of philosophers and social scientists, unequal power dynamics within the household are one of the key causes of existing gender injustice. From a historical perspective, women typically have not had as much bargaining power as their male partners because they tended to either be economically dependent on their male partners as full-time homemakers, or to earn less than their male partners. The imbalance in economic resources and earning potential, in turn, has often led to women’s reduced bargaining power within the household. According to feminist thinkers, efforts to empower women at home, and in particular those who are full-time homemakers and caregivers, should focus on rendering the structure of the labor market more equitable (so as to expand women’s career opportunities), providing families with subsidized childcare, and/or increasing and improving in-cash welfare benefits provided by the state.

Power dynamics at the workplace have long been a key focus of political philosophers working in the Marxist-socialist tradition and are currently of increasing interest to a wider range of moral and political philosophers such as relational egalitarians. This topic
is closely related to the notion of exploitation. Broadly speaking, for a workplace to be fair, its employees must not be exploited (taken unfair advantage of). This means, for example, that the employer must pay the employees a livable wage or a wage commensurate with their experience—even if they would still work for a much lower salary because of lack of other options. One way to incentivize employers to treat workers in a fair manner is to ensure that the workers have enough bargaining power—for example, that they have real exit options, such as another attractive work opportunity. Increasing workers’ bargaining power, then, is a practical route to creating decent and fair working conditions.

In addition to relations between individuals (such as between an employer and an employee or between a husband and a wife), power dynamics and bargaining power are also useful concepts when it comes to theorizing about relations between institutions and individuals governed by, or embedded in, such institutions. Republican political philosophy, for example, emphasizes the value of freedom from dominating control of more powerful agents. Since corporations and public institutions have more economic, legal, and social resources than most individuals, they can in principle easily force individuals to act in certain ways or constrain their options. The expanding philosophical subfields of social ontology and shared agency are likely to make extensive use of the notion of power dynamics in the future.

Further Reading

- [https://plato.stanford.edu/entries/feminist-power/](https://plato.stanford.edu/entries/feminist-power/)
- [https://plato.stanford.edu/entries/contractarianism-contemporary/](https://plato.stanford.edu/entries/contractarianism-contemporary/)
Deontological Ethics

Ethical theory, also referred to as moral philosophy, focuses on what kinds of choices and actions are morally forbidden, required, and permitted. Deontological ethical theories, which are traditionally thought of as inspired by the moral philosophy of Immanuel Kant, are a subset of ethical theories whose starting point is the claim that what makes a given action forbidden, required, or permitted is the intention behind an action and, relatedly, the reason the moral agent had for acting this way. In other words, deontological theories are a subset of ethical theories that are characterized negatively by the fact that they judge the moral status of an action by criteria other than the consequences of the action.

The major alternative to deontology is thus called “consequentialism.” Such consequentialist theories assess the moral status of an action in terms of its consequences. For instance, a consequentialist theory might judge that an action is morally permissible if and only if it brings about the most good (although different consequentialists will have different theories of what counts as “good”). For deontologists, in contrast, some actions are morally forbidden regardless of the consequences. For example, many deontologists will argue that it is wrong to punish an innocent person regardless of how much good would be brought about by doing so. Deontologists will say the same about many actions that most people think of as wrong.

Varieties of Deontological Theories

While deontologists agree in rejecting consequentialism, there are many intramural differences between deontological theories. Three examples of such theories are: agent-centered theories, patient centered theories, and contractarian theories. Agent-centered deontological theories claim that each agent has special obligations and permissions to perform certain actions that other agents lack. For example, agent-centered theories will plausibly claim that a parent has special obligations to look after her child’s welfare and special permissions to prioritize her child over others. Patient-centered deontological theories judge the moral status of an action in terms of the rights of those affected by the action. For example, such theories will deem actions such as murdering the innocent wrong because innocent people have a right to their lives. Another class of deontological theories are contractarian theories. These theories are inspired by the social contract tradition in political philosophy which views political
society as a special kind of contract formed among a group of people. Contractarians apply similar ideas in the moral sphere. For example, one prominent theory in this tradition claims that “An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behaviour that no one could reasonably reject as a basis for informed, unforced, general agreement” (Scanlon 1998, 153).

**Strengths of Deontology**

Deontological theories have many strengths. First, such theories make plausible claims about many actions that most people judge intuitively obligatory or permissible. It is wrong to lie, cheat, steal, kill, and rape. Many people will find such actions wrong even if good consequences follow from them. For example, consider a famous case:

*Doctor:* You are a doctor at a hospital where five of your patients need different organs. Your next patient is a perfectly healthy young man. You can kill him and distribute his organs to your other patients. Moreover, you are certain you can do this without anyone ever finding out and you are certain that this man has no family or friends who will miss him.

Many people would find it abhorrent to kill the innocent man in this case. Deontologists, unlike consequentialists, can agree with this verdict, though different deontologists may have different explanations of why it is wrong to kill the man and harvest his organs.

Deontological theories are also attractive in that they give agents more freedom to pursue their projects than consequentialist theories. For consequentialists, it is wrong of an agent to, say, go to the movies, when she could instead spend her time and money making the world better. Deontologists need be committed to no such thing. Instead, deontologists will typically allow agents to pursue the kind of life they want as long as they don’t violate other people’s rights. Moreover, deontological theories, unlike consequentialist theories, can distinguish between actions that are morally obligatory and those that are supererogatory. These latter actions are those that go “above and beyond” what morality requires of us. If an agent sends half her income to charity or lays down her life to save the rainforest, then she has plausibly done something supererogatory. Consequentialist theories struggle to explain this. The agent who makes great sacrifices to bring about the best consequences merely does her duty, but does not do anything beyond what morality requires.
Relation to Fairness and Challenges

When it comes to the relation between the notion of fairness and deontological ethics, a plausible way of conceiving of fairness would be to treat every moral person in a fair and equal way, thereby respecting their rights and choices in the same way regardless of who the person in question is.

Deontological theories also face challenges. While it is intuitive that we should not punish the innocent to bring about good outcomes, it is unclear how much we should respect this intuition. Deontology can look like a form of “rule worship” as it sometimes requires us to make the world worse for the sake of a rule that deontology deems inviolable.

Further Reading

Consequentialist Ethics

Ethical theory, also referred to as moral philosophy, focuses on what kinds of choices and actions are morally forbidden, required, and permitted. Consequentialist ethical theories first emerged in the writings of British philosophers Jeremy Bentham and John Stuart Mill. They are a subset of ethical theories whose starting point is the claim that what makes a given action good or bad (required or permitted) are the predicted or actual consequences of this action. Consequentialist theories judge the moral status of an action by the consequences of the action. Acts or intentions of a moral agent are thus to be morally assessed solely by the goodness or badness of the states of affairs they bring about. Whichever actions bring about most good out of all the possible actions are those that are morally required and morally right.

Moral Assessment of Outcomes

Consequentialist philosophers differ in the ways they define and specify the “goodness” of various states of affairs. For some, there is only one kind of “the good,” such as desire-satisfaction or pleasure. In the case of classic utilitarianism – a well-known subset of consequentialists endorsed by Bentham and Mill – the only good thing is pleasure and the only bad thing is pain. Consequently, classic utilitarians postulate that we are morally required to act in ways that will bring about most pleasure or happiness for all the individuals involved. For other consequentialists, “the good” can have several different forms (such as pleasure, knowledge, and beauty).

Another crucial component of consequentialist ethical theories is agent-neutrality. This means that the moral assessment of the outcomes of certain actions cannot depend on the particular perspective of a specific agent, but have to be evaluated from the perspective of a by-stander or observer.

Total Consequentialism

What any consequentialist ethical theory must further specify is which individuals’ happiness or pleasure is relevant to the moral assessment of actions, and how the good is to be distributed among the individuals involved. (Typically, non-human animals who are sentient, and can therefore experience pleasure and pain, are included in the moral assessment of actions.) On one particular consequentialist account, what matters
morally is maximizing the total amount of good achieved by an action, regardless of how it is distributed among the individuals involved (total consequentialism). The problems total consequentialism runs into can be illustrated by a well-known thought experiment which constitutes an objection to this type of consequentialism:

*Utility Monster:* There exists a living being—a “utility monster” – who receives much more utility (or pleasure) from each unit of a consumed resource than anybody else on the planet. Even if this resource was divided between everybody else on the planet, it would still collectively bring about less good (pleasure) than for the utility monster. Given that this being can get so much pleasure out of the resources, according to total consequentialism the morally right thing to do is to leave the entire resource in question to the utility monster and give none of it to anybody else.

Total consequentialism focused on maximizing the total positive outcomes of actions has been further criticized due to allowing gross inequalities among individuals. For example, the table below illustrates two ways of distributing resources, in which distribution A and distribution B are equally morally right since they present the same total amount of the resource distributed (12 units):

<table>
<thead>
<tr>
<th></th>
<th>Anna</th>
<th>Carl</th>
<th>Thomas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution A</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Distribution B</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

However, even though distribution A allows the same total amount of resources distributed as B, distribution A is less equal than distribution B. To overcome this challenge to total consequentialism, another kind of consequentialism has been proposed. On this account, what matters morally is maximizing the average amount of good the action has brought about in every individual involved (average consequentialism). This kind of consequentialism values the pleasure of any one individual equally to the pleasure of any other individual and clearly favors distribution B over distribution A in the table above.
Act vs. Rule Consequentialism

Another important distinction within consequentialist theories is between act consequentialism and rule consequentialism. The former postulates that every individual action should be assessed by the actual states of affairs it brought about. Consequently, every person should think about the plausible consequences of the action they want to perform right there and then. By contrast, rule consequentialism postulates that actions should be assessed by generalizations regarding what actions of this type typically bring about. Consequently, people do not have to assess every action they are about to perform, but to learn and abide by general rules of consequentialism. To illustrate the difference between these two kinds of consequentialism, consider the following scenario:

Red Light: Generally speaking, running a red light is a risky and careless behavior, and is prohibited by the law. A person driving a car in a hurry sees that the light in front of the car has just turned red. However, the driver is absolutely certain that there is no one around and that if she runs the red light, no one will be at risk of getting hurt and she will not get a ticket.

According to act consequentialism, running the red light will bring about more value than stopping at the light, therefore the driver would be morally justified to ignore the red light. According to rule consequentialism, however, the driver is required to abide by the general rule “Do not run red lights,” which reflects what typically brings about better states of affairs, even if this is not true in this particular case.

Challenges to Consequentialism

Consequentialist ethical theories at large face numerous challenges. A well-known challenge starts with the observation that many such theories – notably utilitarianism – are impersonal and ignore the distinction between one’s own agency and the agency of other people. Consequently, it does not matter which agent performs a given action. But for any individual, it would make a difference whether, say, she caused someone’s death by pulling the trigger or whether this death was the result of someone else’s actions. The distinction between one’s own actions and other people’s actions is morally important to us. Another, related objection to consequentialist ethics is that it does not allow individuals to pursue their own projects and value distinct things in life, because
everybody is obligated to maximize the common good no matter what causes matter to him or her the most.

Further Reading

Responsibility

Responsibility is a type of relation in which moral agents stand vis-à-vis a certain act or outcome: an agent is responsible for, say, the spread of a virus in their community if the agent caused the spread by one of their actions (e.g., by organizing a massive indoor party without masks) and could have avoided the outcome, perhaps with better deliberation or by gathering more information (e.g., by becoming acquainted with the recommendations of health agencies and acknowledging their expertise). An agent might have somehow contributed to a certain outcome (e.g., the outsourced cooks from the catering company who made the food for the party) but the causal link between their action and the outcome is too intricate or vague to be identifiable.

Likewise, it is typically accepted that a person who is coerced into doing something, or has no control over their own actions due to, say, medical reasons, cannot be said to be responsible for the outcome, or is responsible only in a superficial sense (e.g., an employee who was threatened with dismissal unless they prepared the food for the party; a starving drug addict who steals food from a store). Accordingly, responsibility is more than causation and presupposes a certain degree of freedom in agents.

The notion of responsibility is pervasive in moral theory and philosophers often highlight how it anchors important practices. Insofar as agents are responsible for their actions, they may be blameworthy, held accountable, or praiseworthy. Forgiveness and thankfulness are also attitudes we only hold towards those who are responsible for their actions.

Collective Responsibility

 Debates are still ongoing in philosophy on the conditions for someone (or something) to be a moral agent and, thus, potentially responsible for an action, and on whether actions can really be free in any robust sense. These questions are further complicated when we deal with groups and not individuals. Take the case of a pharmaceutical company whose marketing strategies for certain analgesics cause an opioid crisis. Philosophers disagree whether it makes sense to say that the company is responsible or, in fact, we can only properly refer to the responsibility of individuals: managers, owners, investors, and workers who participated in the decision-making or carried out the
marketing plan. The crucial question is whether “the company” in itself can be said to be a free moral agent that is not reducible to any set of its members.

Usually, a part of the debate on collective responsibility focuses on the procedure of deliberation that took place for a decision to count as the company’s decision: was the procedure hijacked by one or several members of the organization, or by another organization? Were there any inconsistencies in the procedure such that they would make the outcome irrational? Did members interact to make the decision and carry it out? Did they identify themselves as a collective? These questions point to the level of control that the company can be said to have over its own agency and to whether individuals were acting jointly or separately.

If one accepts that there is such thing as the responsibility of a pharmaceutical company, one still needs to answer the question of how responsible each individual in the company is. Common conditions for individual responsibility will apply: for instance, whether they were coerced or unaware of the consequences of their actions. Some level of responsibility, though, may be determined by their position within the hierarchy of the organization and the power that their position gave them.

Recent debates on collective responsibility deal with the responsibility of whole societies in very complex processes that produce injustice, e.g., global supply chains that depend on the extreme exploitation or quasi-enslavement of workers in poor countries and involve consumers all over the world. Some have argued that, in these cases, the question of backward-looking responsibility (i.e., responsibility for actions in the past that make agents blameworthy) might be the wrong focus and we should talk, instead, of forward-looking responsibility, i.e. responsibility for bringing about a state of affairs that is more just (Young, 2011). Although the failure to address an injustice might also make someone blameworthy, forward-looking responsibility is less focused on assigning blame and more centered on compelling agents to fulfill their duties as citizens, e.g., cooperating to make sure every member of their society has the necessary means to live a fulfilling life.

Further Reading


Egoism

In philosophy, egoism is a position within the domain of ethical theory relating to one’s decisions and actions. Broadly speaking, to be an egoist means to act only in one’s own interest and to be motivated exclusively by one’s own wellbeing. In this sense, this term is used in the philosophical context in the same way it is used in the popular, or everyday, context.

However, philosophers have distinguished between several distinct types of egoism that allows them to track more concrete attitudes of practical reason. One of these distinctions is between egoism as a term that describes the way some people behave in the real world and egoism as a normative position – a claim about how people ought to behave.

Psychological Egoism

A well-known descriptive version of egoism is psychological egoism. This type of egoism claims that every human being has only one aim: his wellbeing or welfare. This, more specifically, is the ultimate aim individuals have. Other aims—intermediary aims or means to our ends—can be different (such as actions aimed at other people’s wellbeing), so long as these actions finally lead to an increase in one’s own wellbeing.

The most convincing arguments for psychological egoism come from the observation of people’s behavior and their motivations. People, it is argued, are more frequently motivated by their own self-interest than by ‘altruistic’ reasons. For example, a child might do what she is asked to do faster or better if she is promised candy once the chore is completed. Another set of arguments in support of psychological egoism comes from evolutionary theory. Since evolution favors individuals who prioritize their own wellbeing over the wellbeing of other members of their group, the argument goes, it is likely that all, or at least the majority of, people are psychological egoists. But there are good objections to this position. One obvious objection is that while it is true that individuals often, or sometimes, act in their own interest, it is deeply implausible that they are always ultimately aiming at maximizing their own wellbeing. For instance, an action of a brave soldier who sacrifices his life while defending his nation can hardly be described as aiming at maximizing his own wellbeing without stretching the idea of what it means to aim at one’s own wellbeing.
Another objection to psychological egoism may come from evolutionary theory itself: according to the selfish gene hypothesis, it is our genes and not our whole bodies that aim at survival. Consequently, we are biologically determined or ‘programmed’ not to act in our own self-interest understood as maximizing our own welfare, but rather in the interest of those biologically related to us. But acting so as to maximize the welfare of our biological family is not the same as acting in our own, narrowly conceived, self-interest. A third objection to psychological egoism worth noting starts from the observation that people who believe psychological egoism is true conflate the idea of aiming at one’s own welfare and the idea of being motivated by one’s own desires and reasons. The thought that we always act by being motivated by our own desires and reasons is trivially true—because my actions and decisions are mine, I cannot be motivated by anything other than my own desires and reasons, metaphysically speaking. But this does not mean that these desires are always self-regarding. I can be motivated by my own desire to increase the wellbeing of another person, and the fact that this is often the case provides compelling evidence against psychological egoism.

Ethical Egoism

One well-known version of normative egoism—the position that people always ought to act in ways that maximize their wellbeing—is ethical egoism. Ethical egoism is the claim that an individual morally ought to perform a specific action if and only if this action maximizes this individual’s wellbeing, and because it does so. According to ethical egoism, more specifically, what is morally required of me is that I always act in ways that maximize my own wellbeing. Helping someone or something else is morally required if and only if (and because) doing so would ultimately maximize my own wellbeing; likewise, I should refrain from harming someone else only when, and because, doing so is for my own benefit.

Importantly, whether ethical egoism is a correct normative framework does not bear on the truth or falsity of the descriptive claim that people always act in their own self-interest—of psychological egoism.

One possible objection to ethical egoism is that this position is self-contradictory or incoherent. This objection goes as follows. To be an ethical egoist, I must believe that maximizing my own wellbeing is morally required. But if it is morally required, then it is a
requirement not only for me, but for other moral agents too. Consequently, I must believe that other persons are morally required to maximize my wellbeing, and not their own; hence, that others cannot be ethical egoists.

This objection has been challenged in several ways. For example, it has been argued that what ethical egoism may prescribe for me to do (and what is morally required in my case) does not have to be morally required for everyone, or “good absolutely.” If this is true, then, arguably, ethical egoism is not an incoherent position. Nonetheless, few philosophers regard ethical egoism as a serious moral position (often questioning whether it is a moral position at all), and few people are guided by this normative position in their life.

Further Reading

- [https://plato.stanford.edu/entries/egoism/](https://plato.stanford.edu/entries/egoism/)
- [https://1000wordphilosophy.com/2020/02/02/ethical-egoism/](https://1000wordphilosophy.com/2020/02/02/ethical-egoism/)
Freedom and Liberty

The terms “freedom” and “liberty” are used in a variety of ways within philosophy. In metaphysics—a branch of philosophy dealing with questions about the nature of reality and existence—the discussions of freedom typically take place through questions about the possibility of free will, the compatibility of human freedom and biological or physical determinism, and the different types of causal relations that occur between different agents and objects. In ethics, freedom is frequently discussed in the context of moral responsibility for one’s actions and in the context of moral blameworthiness. Most ethicists would agree, for example, that an agent who did not pull the trigger of a gun “freely,” but was coerced to do so by another party (either by physical force or by a verbal threat), cannot be morally responsible for the damage the gun has caused, and hence cannot be morally blamed for it, even if he is causally responsible for the gunshot.

This entry, however, will focus on the concept of freedom or liberty as they are discussed in political philosophy, not in metaphysics or ethics. For simplicity’s sake, these two terms (freedom and liberty) will be used interchangeably in this entry. The long-standing tradition of political thought, especially in modernity and the contemporary world, can be framed as a struggle to reconcile the values of liberty and equality in various ways.

Utilitarian Framework of Liberty

The utilitarian framework of political thought, which dates back to Jeremy Bentham and John Stuart Mill, prioritizes equality over liberty and postulates an equal treatment of every human being (or even every living being). However, Mill can also be credited for coining the “Harm Principle,” which has played a great role in political thought ever since, and according to which every individual should be permitted to act freely and choose a lifestyle for herself so long as she does not thereby harm another individual, or impede on the liberty of another human being. Mill’s treatment of the value of liberty is thus intertwined with the value of equality, since every individual’s liberty is supposed to be comparable to that of other people.
Liberal Egalitarianism

Another important tradition of political thought, liberal egalitarianism (most commonly associated with the work of John Rawls), is also concerned with reconciling the value of equality with the value of liberty. The liberal egalitarian emphasis on liberty can be understood as striving to ensure that everybody has the freedom to develop and pursue their own conception of the good life. The value of equality is also essential and corresponds to the liberal commitment to supporting individual choices regardless of their preferences. Such respect is considered to be incompatible with very large socio-economic inequalities. Rawls has famously unified these values in his two fundamental Principles of Justice. The First Principle gives everyone equal right to a scheme of basic liberties compatible with the same liberties for others. The Second Principle specifies that socio-economic inequalities must be arranged to everyone’s advantage and that everyone must compete for jobs and other positions on equal terms.

Libertarian Framework

By contrast to Mill or Rawls, the libertarian tradition of political thought (commonly associated with Robert Nozick, Milton Friedman, and Friedrich Hayek), explicitly prioritizes liberty over equality. Libertarians are primarily concerned with providing everybody, including the least well-off, with a large degree of freedom from external constraints (such as scarcity of basic resources or an abusive work contract). In addition, some libertarian philosophers emphasize the normative priority of individual freedom and advocate for a small state with minimal intervention in private and market-based interactions. According to Milton Friedman, for example, individual freedom is protected by maintaining a decentralized market which is separate as much as possible from politics and the coercive apparatus of the state. The sole function of such a state would be the specification, interpretation, and enforcement of negative individual rights through the military, police, and court system.

Negative vs. Positive Liberty

A further important distinction within the idea of freedom in the context of political philosophy is that between negative liberty and positive liberty, most prominently discussed by Isaiah Berlin in his essay “Two Concepts of Liberty.” Berlin characterizes negative liberty as the absence of constraints on, or interference with, an agent’s possible actions. This concept has been taken up by republican tradition in political
philosophy (represented, for example, by Philip Pettit). Republican philosophers—which should not be confused with the Republican Party in contemporary U.S. politics—are concerned with the idea of freedom from domination: the absence of dominating control or the condition of not being subject to the arbitrary or uncontrolled power of others. They understand the idea of dominating control as a hierarchical mechanism whereby one individual or institution can exercise arbitrary power and control over another individual. Positive liberty, by contrast, is characterized by Berlin as the capacity to determine oneself or to be in control of one’s destiny, for which it is necessary that one is aware of multiple options and choices regarding ways of living. A crucial way of ensuring that individuals attain a high degree of positive liberty is by providing them with appropriate schooling.

Finally, liberty continues to play an enormous role in the formation and maintenance of liberal democratic nation-states. In this context, specific liberties are often discussed and embedded in countries' constitutions and amendments, such as the civil liberties protected by the first Ten Amendments to the U.S. Constitution. Such civil liberties embedded in fundamental national and international laws and regulations have been largely inspired by the work of political philosophers, notably Immanuel Kant, Mill, and Rawls.

Further Reading

- https://plato.stanford.edu/entries/liberty-positive-negative/
- https://plato.stanford.edu/entries/utilitarianism-history/
- https://plato.stanford.edu/entries/liberalism/
- https://plato.stanford.edu/entries/libertarianism/
- https://plato.stanford.edu/entries/republicanism/
Cooperation

Cooperation is a collective activity between two or more agents in which they all extract mutual benefits from each other’s efforts. It is a ubiquitous phenomenon in social life and is at the core of two fundamental questions in political philosophy: how to ensure that people cooperate to guarantee common goods and how to allocate the benefits and burdens of cooperation. Outside political philosophy, the literature in philosophy of action and social ontology has focused on determining the conditions for a joint activity to count as cooperative.

Definitions of cooperation may differ on the type of knowledge, intentions, and reasons that agents are supposed to have for their collective activity to count as cooperative. It is commonly accepted, though, that some level of coordination is required but not sufficient. The agents’ actions must have a more robust relation, which has been conceptualized in many different terms. For instance, Michael Bratman (1992) has argued that the intentions of cooperative agents need to be adaptable to and compatible with each other, and agents must be mutually supportive.

The Assurance Problem

Traditionally, political philosophy has focused less on definitional questions about cooperation and more on what are commonly called the “assurance problem” and “prisoner’s dilemmas,” which are at the core of any theory of the state. The assurance problem refers to the fact that, in order to cooperate, individuals need to be assured that others will do their share. Prisoner dilemmas are situations where for each individual it is rational (i.e., best for their self-interest) to defect when everyone else is cooperating, and yet universal defection is worse for everyone than universal cooperation. Political philosophers have developed theories to address both problems in the organization of society, although often implicitly (both problems weren’t properly conceptualized until the 20th century). Plato, for example, thought cooperation could be ensured by indoctrinating citizens, while Hobbes argued that the solution consisted in sanctions coordinated by a central agent, since the threat of punishment would deter people from defecting (see Weithman, 2010).

Contemporary theories of justice and equality have been more explicit about the centrality of the idea of cooperation. John Rawls (2001), for example, defines society as a
“fair system of cooperation,” meaning that the benefits and burdens of collective activities ought to be distributed fairly, and the rules and procedures that guide cooperation are publicly recognized and accepted as appropriate by citizens (6). Likewise, Elizabeth Anderson (1999) has applied the term of “fair system of cooperation” to the type of economy required by democratic equality. For Anderson, the economy is cooperative insofar as its products are jointly produced by everyone working together and there is a just division of labor and of the fruits of labor.

Morality and Cooperation

These theories also draw on the fact that, although strong moral commitments may not be necessary for cooperation, some morally objectionable situations are incompatible with cooperation. For instance, it is implausible to claim that a thief and their victim are cooperating, since the victim is being coerced and being left worse off by giving the thief what they want. This is why cooperation precludes exploitation. If a group in society extracts disproportionate benefits from others’ efforts, then they are no longer cooperating.

Cooperation is also at odds with competition. When agents compete against each other, either one or the other obtains the goods at stake. Competition is a zero-sum game, even when it requires coordination between agents, whereas cooperation is supposed to benefit everybody (although perhaps unequally).

Further Reading

Reasonableness

In a colloquial or general context, the term “reasonable” is used to refer to a person who is agreeable, willing to listen to the rules of reason, and eager to consider to what others have to say; it can also be used to refer to an idea or proposal that exhibits similar features. However, in political philosophy, “reasonableness” is a technical term with a narrower (albeit sometimes contested) meaning. The concept of reasonableness is a key concept in the context of liberal political theory – a framework grounded in the acceptance of various ways of living without privileging some over others, so long as one individual does not harm another one.

Reasonableness is especially essential to John Rawls’s liberal political theory developed between the 1970’s and the 2000’s. For Rawls, reasonableness is a normative standard of thinking or deliberating that is distinct from rationality. To be rational, broadly speaking, means to think in a coherent and logical way. Additionally, for Rawls, rationality is synonymous with being able to decide what one wants (from high-stakes matters such as choosing a career to minor issues such as deciding what to have for dinner) and being able to determine what means to take in order to achieve one’s goals. Rationality therefore has to do with questions of prudence—of doing things for one’s own benefit.

By contrast, reasonableness does not have to do with prudential issues, but rather with moral and political ones where one is motivated by a concern for the wellbeing of others. Reasonableness, then, concerns our moral relationships with other people and the establishment and preservation of a just and stable liberal polity. More specifically, in his early work as well as his later work Rawls claims that for a person to be reasonable means for her to have a “sense of justice.” This, in turn, means to have a moral capacity to judge matters as just or unjust, to desire to act accordingly, and to have a moral motivation to do what the rules of justice require; this standpoint is incompatible with egoism. To be reasonable in this sense is a normative ideal that every citizen of a society should aspire to.

In Rawls’s later work (2005), the adjective “reasonable” also describes a set of beliefs that a person holds. A set of beliefs (or a “comprehensive doctrine” in Rawls’s language) may be described as reasonable when it characterizes and organizes values in a coherent and
logical way; when it draws long-standing upon a tradition of thought; and when it acknowledges the existence of other reasonable doctrines. In a just liberal society, Rawls claims, citizens exhibit different reasonable comprehensive doctrines (which depend on their upbringing, among other things), which is a productive state of affairs as long as unreasonable doctrines do not dominate the political spectrum. Rawls’s later work also applies the adjective “reasonable a set of principles of justice—specifically, those proposed by Rawls himself. On this account, reasonable principles of justice are characterized by the feature of publicity: they are known and acceptable to all relevant members of a community or nation-state. Here, reasonableness is closely related to the concept of “public reason,” widely discussed in contemporary liberal political theory by thinkers such as Gerald Gaus, Jürgen Habermas, and Rawls himself. The idea of public reason concerns the way we choose fundamental socio-political principles that regulate our coexistence in a society. It starts from the conception of persons as free and equal beings. Public reason requires that political rules that govern our common life be justifiable or acceptable to all those individuals over whom the rules have authority.

Further Reading

Diversity

Diversity refers to the range of differences among the members of a certain set with respect to some feature (e.g., race, gender, age, etc.). When it comes to diversity among groups of human beings, debates in political and social philosophy focus on whether diversity has intrinsic or instrumental value: is it good in itself, or morally required, to have racially or gender-diverse representative bodies, or is diversity good if and only if it produces certain outcomes, e.g., better policies for race or gender minorities?

The answer to this type of question will depend partly on the environment or institution that is being examined. For instance, insofar as political representation in a democracy is supposed to guarantee equality among citizens, it is often argued that democratic equality requires democratic institutions like Congress, town halls, or public agencies to reflect the most salient differences among citizens. This is called “descriptive representation,” i.e. a type of political representation where the representatives resemble the represented and, in addition, the composition of representative bodies matches the distribution of certain features in the general population (although proportionality might not be strictly required).

In contrast, the very nature of an institution might make diversity, or certain types of diversity, unnecessary. For instance, a supporting group for LGBTQI+ individuals might actually fail in fulfilling its role if it is required to have a majority of heterosexual, cis-gendered individuals that matches the proportion of this group in the general population. This kind of “diversity” would be contrary to the group’s egalitarian goal, taken to be legitimate. The situation is different when groups or institutions actively seek to generate unjust inequalities. It would be pointless to raise the question of diversity in the Ku Klux Klan, for example. More ambiguous cases include private institutions where it is still controversial whether the state should mandate diversity, e.g., should a prestigious golf club be required to have racially or gender diverse members?

Political Argument for Diversity

There are at least three different arguments for diversity in public and private institutions. A political argument, which applies to institutions with a significant impact on society, is that a diverse body of representatives is more likely to recognize, understand, and respond to the needs of the groups they represent (Williams, 1998;
Young, 2000). This responsiveness, of course, is conditioned on other factors, like the other social groups to which representatives belong, or the quality of the institutions themselves. Although a Black congresswoman might be more likely to understand the experience of Black women, her socioeconomic background or the influence of lobbies in Congress might impede her from advocating for policies that benefit poor Black women.

**Epistemic Argument for Diversity**

An epistemic argument is that a diverse group of people is more likely to make better decisions. Diversity favors the exchange of different points of views that could make important contributions to the deliberation and allows for parties to point out each other’s blind spots. This argument might apply to political institutions as well as firms or schools. A crucial question here is what kind of diversity really provides better epistemic results, and whether cognitive diversity is the one that matters primarily instead of, say, diversity in values or socio-economic backgrounds (see Landemore 2013).

**Symbolic Argument for Diversity**

Further, a symbolic argument, often made in the public sphere, claims that diversity in positions of power and visibility have positive psychological impact among minorities. Such positions create aspirational models and influence social norms and values. If most models on the cover of magazines are White, for instance, then the standards of beauty will be modeled after the White phenotype and assimilated by most of the population, including people of color, who will lack recognition in their own society. This generates great psychological costs for them. Likewise, as long as there are no women presidents in the United States, women’s trust on their own democratic institutions as well as the authority of the latter might be undermined, while the endeavor of gaining representation will still look (as indeed is) hard to surmount. This might be discouraging for many women.

Finally, it is worth noting that diversity might also be a feature of things other than institutions themselves. For instance, when designing a syllabus, a Literature professor might try to make it as diverse as possible by including authors that are usually underrepresented in the literary canon. This action might give students a better sense of the material and make them think about questions that are often excluded from mainstream discussion and are nevertheless important.
Further Reading

Multiculturalism

Many people understand the term “multiculturalism” as a feature about contemporary Western societies—multiculturalism denotes the fact that such societies are composed of individuals whose cultures, religions, and languages differ from one another. While this meaning is helpful for philosophical theorizing about the fact of diversity, when socio-political philosophers discuss multiculturalism they typically focus on its normative (evaluative) dimension. Normative theorizing about multiculturalism involves questions such as: Are multicultural countries more desirable than homogenous ones? What are the advantages of multicultural citizenship? Should cultural minorities be expected to assimilate to the dominant culture around them, or should provisions be made to accommodate the minority cultures? To what extent should societies adapt and change to accommodate the cultures, ways of living, and languages of minorities?

Broadly speaking, philosophers who advocate for multiculturalism make the normative claim that countries characterized by cultural diversity should not suppress, ignore, or even merely tolerate minority practices, but rather accommodate and recognize them within the public domain. One example of such a treatment would be the recognition of the language spoken by a minority culture as official by the state government. In this sense, multiculturalism is closely connected to other movements for the inclusion of marginalized social groups, such as the movements for LGBTQ rights or for the rights of individuals with disabilities.

Accommodating diverse cultural practices might require discharging so-called “group-differentiated rights”: specific exemptions or provisions which, while not required by members of the dominant culture, enable members of the minority culture fully to participate in the civic life of the society without compromising their beliefs and ways of life. These rights include exemptions from federal or local laws, provision of free translation services in official contexts, and affirmative action policies. In addition, multiculturalist societies should strive to acknowledge and represent minority cultures and practices in the public domain. They can, for example, include a minority’s historical perspective in the educational curriculum, fund museum exhibitions, and ensure adequate representation of the interests of cultural minorities in government bodies.

There are several philosophical justifications of the claim that cultural minorities should be recognized and accommodated within the socio-political sphere. One such
justification centers around the notion of equality. The society as a whole is responsible for remedying the unchosen disadvantages stemming from a minority status because, first, members of minority groups did not (typically) choose to be in a socially disadvantaged position, and, second, cultural membership is a significant part (if not a necessary condition) of shaping one’s character, identity, and understanding one’s options. Another philosophical justification for multiculturalism stems from the idea that freedom requires not being dominated or limited by other agents and forces beyond one’s control. According to this line of thought, culturally separated and divided societies may overlook that certain individuals from minority cultures might not have the opportunities to flourish and might not be presented with a wide range of life choices. (This, of course, can characterize the dominant culture as well.) Multiculturalism can thus be a way of resolving power dynamics within cultural groups and enabling minority individuals with more freedom. According to yet another justification, multiculturalism is a way of (partially) rectifying historical injustices such as the suppression of cultural practices, religions, and languages of minority groups.

There are many philosophical critiques and nuanced discussions of the notion of multiculturalism. One example of such criticism is the claim that multiculturalism as a movement opposed to assimilation and integration might cause separation, division, and lack of mutual understanding among neighboring groups of people. Another line of criticism challenges the methodology of multiculturalism itself, claiming that the multiculturalist movement is rooted in a liberal framework and thus is opposed to non-liberal cultural practices.

Further Reading

- [https://plato.stanford.edu/entries/multiculturalism/](https://plato.stanford.edu/entries/multiculturalism/)
Representation

Representation is a type of relation in which an agent (the representative) speaks, decides, and acts in the name of another agent (the represented). In virtue of this relation, the latter becomes present (re-presented) through the former. For example, although 67 million British people cannot be literally present at a NATO Summit, their voice is made present in some sense through their prime minister.

We can find instances of representation in many different spheres: between parents and children, CEOs and companies, NGOs and oppressed groups, lawyers and clients, etc. Political representation can take at least three forms. As it is the case in authoritarian regimes, a representative may claim to act "in the name" of their citizens (for instance, by declaring war on another country) without any intention to track their interests or preferences. To the contrary, representatives in a democracy are supposed to be either delegates or trustees. These two ideals of democratic representation go back, respectively, to the writings of James Madison and Edmund Burke in the late eighteenth century. For Madison, government is delegated to a small number of people who aggregate and adjust the clashing views of the public. Representatives are not always enlightened and, thus, may fail to see beyond their party's interests. In the expression of preferences, a large and diverse population gives representatives some cues to mitigate the effects of those failures. Instead, Burke considered that representatives are best thought of as trustees because people often fail to identify their own interests. Here representatives must use their judgment to identify such interests and act accordingly, regardless of what the represented express.

Delegation and trusteeship are, in practice, complementary. Citizens expect representatives to advance their interests and exercise good judgment in issues where citizens themselves lack proper knowledge or where it would be impractical to carry out elections. More broadly, it is very likely that, through propaganda, advocacy, or their own policymaking, representatives end up shaping their constituents' preferences, in which case we cannot truly say that citizens' preferences exist independently of the representative's judgments and actions.
Theories of Representation

An important question in theories of representation is how representatives gain their authority, that is, their moral power to act in the name of citizens. Most theories have focused on formal representation, which has its origins in and is exercised through institutional mechanisms of election, selection, and decision-making. In such cases, the authority of representatives and their actions stems from procedures deemed to be fair, like a democratic election. But if such procedures are absent, then representation may be informal. Take the case of the Occupy Wall Street movement, who claimed to "be" the bottom 99% of the income distribution even if the "99%" did not elect them. In this type of representation, the fact that an agent has advocated for the interests of a group, despite lacking formal authorization from that group, might sometimes be enough to call the agent a representative.

 Democracies, especially when there are persistent and marginalized minorities, often face the demand for representatives (both formal and informal) to resemble the represented, in which case there is descriptive representation. As Hannah Pitkin notes in *The Concept of Representation* (1967), the emphasis here is rather on the composition of representative bodies, not necessarily on what they do. This is why the push for descriptive representation often leads to adopting a principle of proportionality in the structure of decision-making bodies so these can reflect the diversity of the population. Defenders of descriptive representation argue that sharing physical features, as well as cultural or socioeconomic backgrounds with the represented, may have important effects on the self-respect and visibilization of minorities. Further, it makes it more likely for representatives to know and act in favor of the interests of the represented. Critics, instead, worry about the accountability of descriptive representatives insofar as their actions seem to lose importance in the justification of their status as representatives.

Although democracy requires equal representation, this does not mean that every interest can or should be equally represented. Apart from the fact that it is unviable for political institutions to consistently accommodate every possible claim to representation, some claims will have to be deliberately excluded according to the principles of legitimacy that such institutions ought to honor. Now, what criteria should be employed for this exclusion? Are there interests that ought to be prioritized? A minimum and intuitive requirement for representation is that claims fall within the scope of what is reasonable, that is, of what is justifiable to other citizens without being
imposed on them (see entry on Reasonableness). But reasonableness alone won’t take us too far, since it only solves unambiguously the most obvious cases (e.g. the claims of the members of the Ku Klux Klan). More recently, some authors like Melissa Williams (1998), Iris Marion Young (2000), and Suzanne Dovi (2009) have defended the idea that the oppression a group has endured or continues to endure is the most relevant factor to legitimize its claims to political representation and prioritize their access to representation. Initiatives like the reservation of seats for certain minorities, intraparty quotas for women, or racial districting are rough applications of this “oppression principle.” A corollary for the principle, according to Dovi, is that democracies ought to marginalize two types of groups from representative bodies: those who oppress and those whose privileged status helps to maintain the oppression.
Transparency

Transparency is a requirement to reduce or eliminate obstacles that render information more visible or accessible. Typically, this is a requirement on collective and individual agents who hold information that may be relevant for other agents’ decision-making. Although the issue of transparency is often raised with respect to technology companies and their complex algorithms, the demand for transparency can be made to any organization, public or private, that is to be held accountable for its decisions.

Quantity vs. Understandability

Philosophical debates on transparency deal with the question of whether, on the one hand, it is the amount and type of information that matters for an organization to be transparent or, on the other, there is also a duty to make the information understandable for the public or to explain how the information was produced. Furthermore, it may be that transparency is not an end in itself. For instance, Turilli and Floridi (2009) have argued that transparency is not a principle in itself but, instead, a condition that enables an organization to honor ethical principles, such as accountability, anonymity, safety, or privacy.

Types of Transparency of Algorithms

Within the literature on algorithms, Creel (2020) has helpfully distinguished three types of transparency: 1) functional transparency, which allows users to know which algorithm a certain program instantiates, that is, one can know the rules by which certain inputs are transformed in certain outputs; 2) structural transparency, which provides knowledge of how the code as written (the same algorithm may be realized in multiple codes) produces the result of the program; 3) and run transparency, by which one can know the program that was run in a particular instance, including the specific machine and input data that were used. These forms of transparency are independent from one another: for instance, one can know which algorithm was used without knowing how the code instantiates it or which hardware or data were used.
Further Reading

- Vredenburgh, K. The right to explanation. Journal of Political Philosophy, n/a(n/a). https://doi.org/10.1111/jopp.12262
Accountability

Accountability is a requirement on representatives to justify their actions to the represented and be subject to sanctions (or even removal) if necessary.

Two crucial questions in political philosophy are when and how representatives are to be held accountable. The former line of inquiry deals with the kind of actions that should trigger accountability processes and, more specifically, whether representatives should be held accountable when they fail to satisfy the preferences of the represented or do not protect the interests of the represented. This dichotomy, however, has undergone criticisms that point to the fact that representatives may also shape the preferences of the represented (e.g., by campaigning on the successful policies of their term), so people’s preferences are not independent from their representatives’ actions.

The question of how accountability should be ensured refers to the mechanisms that make it more likely for representatives to fulfill their duties. Elections are often taken as the most important device in this respect. However, recent philosophical and empirical research (Mansbridge, 2004; Achen and Bartels, 2017) has highlighted the limitations of electoral processes in guaranteeing accountability.

Accountability and Representation

Political philosophers also debate whether certain types of representation are particularly vulnerable to accountability deficits. For instance, it is often argued that whenever representation is based on a resemblance between the representative and the represented (descriptive representation), accountability is likely to be weak. In such cases, representatives may seek to strengthen the resemblance without being responsive to the demands of the represented. Likewise, the represented may be forgiving by virtue of feeling themselves “reflected” on the represented, even when the latter is performing poorly.

Some theorists of political representation have thought of accountability as the paradigmatic feature of a representative, in contrast with those who argue that authority is what characterizes representation (Pitkin, 1967). The authority view claims that a representative is someone who has been authorized to act, that is, has been granted the moral power by the represented to decide and act on their behalf and to request or
forbid actions to the represented themselves. In contrast, the accountability view argues that a representative is someone who must answer to other people for what she does, that is, who is to be held accountable. While in the authorization the representative is somehow freed from responsibility because she has been authorized to act as such, representatives in the accountability view are the ones who are responsible to the represented.

Both views have been criticized and it is now more plausible to say that representation often involves both authority and accountability, among other elements. However, both authority and accountability may look very different depending on whether it is a case of formal or informal representation (where representatives haven’t been selected through systematized election or selection procedures, see Salkin, 2021). Consider the case of a celebrity speaking on Twitter in the name of victims of sexual abuse, being a victim herself. No one has appointed her to fulfill that role, so the source of her authority seems to be her own experience, but this is open to debate. Moreover, the only channel for the represented to hold her accountable may be social media (e.g., direct messages, unfollowing, viral campaigns against the celebrity, etc.), but it is unclear how effective these mechanisms could be in ensuring proper representation.

Further Reading

Envy Freeness

Envy-freeness is a criterion of fair division of goods. It states that the allocation of a resource is fair as long as every party with equal rights is assured that she is receiving a share that is as good as the others’ share, so no one prefers someone else’s bundle to their own. This criterion has been discussed mostly from a game-theoretical approach in economics. It does not imply, though, that agents actually experience envy as an emotion, characterized by some level of hostility towards the envied. Instead, the criterion only says that whenever someone prefers someone else’s bundle to her own, then the distribution is not fair (D’Arms, 2016).

In political philosophy, Ronald Dworkin (1981) placed the envy-freeness criterion at the center of his theory of equality of resources. In a thought experiment, Dworkin imagined a market where everyone receives the same amount of resources to buy what they want. At this point, the no-envy test is met because no one can prefer no one else’s bundle insofar as they are the same. But once production and exchange start, the distribution is no longer envy-free: people make gambles that leave them with a bundle they may deem worse than others. Now, Dworkin distinguished between two types of gambles: those where the risks are accepted by the agent and could have been declined (e.g., deciding to use one’s bundle to buy a very expensive bottle of wine and then find oneself without money to pay the rent), and those where the risks were not deliberately accepted (e.g., one buys a house that, due to a recession years later, depreciates significantly). For Dworkin, only the latter warrants a redistribution of resources. He also considered that the envy test should be applied throughout people’s lives (and not just when they are born, as his thought experiment would suggest) to make sure that equality of resources is guaranteed all along.

A larger debate in political philosophy goes beyond the envy-freeness criterion as it is adopted by Dworkin or by economists and focuses on envy as a psychological motive in concerns about equality. Most notably, John Rawls (1999, 464-68) excluded envy (taken as an actual emotion) from the considerations to determine whether a distribution is just. Rawls viewed envy as a vice that is excusable only when it is a response to an actual injustice and remains within reasonable limits. Envy becomes inexcusable when it is a response to a justified inequality. It is an emotion that does not stop at the mere realization that someone else’s bundle is better than one’s own: it tends to fill the
envious person with hostility against others. She would be willing to deprive them, even when such deprivation would make herself worse off, just for the sake of reducing the inequality. Further, envy has costs for society at large: it seeds distrust, since the awareness that someone is envious might lead people to distance and protect herself from the envy person. Nevertheless, Rawls acknowledges that if the inequalities sanctioned by the principles of justice produce so much excusable envy, then the principles might have to be reformulated.

Further Reading