Issues in Ethics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Introduction</th>
<th>01:: 'The Moral Status of Terrorism' from <em>The Ethics of War and Peace, 2nd Edition</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>02:: 'Duties of Humanity' from <em>The Ethics of Global Poverty</em></td>
</tr>
<tr>
<td></td>
<td>03:: 'Holding Psychopaths Responsible and the Guise of the Good' from <em>Current Controversies in Bioethics</em></td>
</tr>
<tr>
<td></td>
<td>04:: 'Diachronic Identity and the Moral Self' from <em>The Routledge Handbook of the Philosophy of the Social Mind</em></td>
</tr>
<tr>
<td></td>
<td>05:: 'Human Genetic Engineering' from <em>Engaging Bioethics</em></td>
</tr>
<tr>
<td></td>
<td>06:: 'The Basic Argument for Vegetarianism' from <em>The Animal Ethics Reader, 3rd Edition</em></td>
</tr>
<tr>
<td></td>
<td>07:: 'Solidarity near the End of Life' from <em>Ethics at the End of Life</em></td>
</tr>
</tbody>
</table>
Visit our website to discover more

Visit www.routledge.com to order your copy of these titles and get 20% off using discount code IIE17. This code is valid until 31st December 2017 and cannot be used in conjunction with any other offer.
Introduction

This FreeBook brings together a selection of chapters from across our range of publishing in the field of Ethics.

The Moral Status of Terrorism is taken from Helen Frowe's textbook, The Ethics of War and Peace, 2nd Edition. Looking at conceptual and moral distinctions of terrorism, it asks whether terrorism can ever be morally justified. The second chapter, Duties of Humanity is taken from The Ethics of Global Poverty, the second title in our Ethics of.. textbook series. In this chapter, author Scott Wisor discusses the limits of our moral responsibilities to the needs of others.

Chapter three, Reflections Concerning Moral Objectivity by Ruth Anna Putnam, is taken from Facts and Values edited by Giancarlo Marchetti and Sarin Marchetti. In this essay, the author argues that facts and value are equally objective. Next, we offer a chapter from The Routledge Handbook of the Philosophy of the Social Mind. In Diachronic Identity and the Moral Self, Jesse J. Prinz and Shaun Nichols theorise that moral continuity is central to ordinary beliefs about what makes someone qualify as the same person as they advance through life.

Engaging Bioethics is the new textbook from Gary Seay and Susana Nuccetelli. In Human Genetic Engineering they discuss the moral problems around prospective human cloning and the altering of genes for human enhancement. Chapter five, The Basic Argument for Vegetarianism is taken from the third edition of The Animal Ethics Reader. In this essay, James Rachels discusses Peter Singer's argument that it is wrong to cause pain unless there is a good enough reason to cause pain, and asks how this affects or fails to affect people faced with the question of choosing vegetarianism.

Finally, Solidarity near the End of Life is taken from Ethics at the End of Life edited by John Davis. This chapter by Bruce Jennings explores the role of the concept of solidarity and closely related notions in the ethics of end-of-life care.

Visit our website to view information on the books in full, or to purchase a copy. Links are provided at the beginning of each chapter of this FreeBook. If you have any questions, please contact us.

Note to readers: References from the original chapters have not been included in this text. For a fully-referenced version of each chapter, including pedagogy, footnotes, bibliographies, references and endnotes, please see the published title. Links to purchase each specific title can be found on the first page of each chapter.

As you read through this FreeBook, you will notice that some excerpts reference previous chapters – please note that these are references to the original text and not the Freebook.
The Moral Status of Terrorism
The recent surge of interest in just war theory is attributable in no small part to the significant and prolific rise in terrorist activity over the past decade. Of course, terrorism is not a new phenomenon. The term ‘terrorism’ was first coined during the French Revolution in the eighteenth century, and terrorist activities go back as far as Ancient Rome. The IRA (Irish Republican Army) employed a sustained campaign of terrorism against the British in the 1970s, and ETA (Euskadi Ta Askatasuna) has been blowing things up in Spain for the past forty years. What has changed lately is the amount of interstate terrorism. In the past, terrorism was a largely domestic problem, with small groups using violence against their own governments or occupying forces in a bid to change domestic policies. Now, terrorism has an undeniably international dimension, crossing borders in a way that seems to take it beyond the realm of ‘ordinary’ criminal violence and into the realm of warfare. These international terrorists primarily seek to change not the domestic policies of their own countries, but the foreign policies of the countries they target. In this chapter we shall look at whether there is anything conceptually distinctive about terrorism, and whether there is anything morally distinctive about terrorism. We shall also consider whether terrorism can ever be morally justified.

**WHAT IS TERRORISM?**

It is unlikely that there is any single feature (or group of features) that is found in all and only terrorist acts. For example, some people characterise terrorism as violence perpetrated by non-state actors – that is, by people who do not represent the state. But we can think of countless acts of violence perpetrated by non-state actors that are clearly not terrorism. A pub brawl is violence perpetrated by non-state actors, but it is not terrorism. Moreover, some writers argue that there are actions that are terrorist but not violent, such as failed terrorist attacks and terrorist threats.

So rather than try to identify a single defining feature of terrorism, we shall look at a group of features that are commonly thought to be relevant to assessing whether an action counts as terrorism:

1. It is politically motivated.
2. It is aimed at non-combatants.
3. It is perpetrated by non-state actors.
4. It is violent.
5. It is intended to generate fear.

Some instances of terrorism will display all of these features, but others may display only some. We shall consider each feature in turn.
Political motivation

That terrorism is politically motivated is perhaps the least contested aspect of terrorism. It is this feature that is usually thought to distinguish terrorists from a common or garden mass murderer or serial killer. Jack the Ripper may have killed innocent people and spread terror amongst the population of London, but not many people think that the Ripper was a terrorist. The most obvious difference between the Ripper and, say, Al-Qaeda, is that the Ripper’s killings were not politically motivated. Unlike ordinary murderers, terrorists have a political agenda that their attacks are intended to further.

This agenda may be to induce a state to cease the occupation or invasion of another state, to ban animal experimentation, to halt the development of nuclear weapons, or to reduce environmental damage. Whatever its content, the agenda matters because it is this that gives terrorist attacks their instrumental (or, as it is sometimes called, their symbolic) nature. Terrorist attacks are not ends in themselves; they are part of a broader plan. For example, animal rights terrorists do not try to stop animal experimentation by systematically killing all scientists involved in such experiments. Rather, they kill a few scientists in order to scare other scientists into halting their experiments, and to induce citizens in general to pressurise their government into banning animal experimentation. The effectiveness of the initial killings in achieving the terrorists’ goal comes from preventing the targeted scientists from carrying out experiments, but from the wider political effects of the killings. Ultimately, terrorist attacks are not about killing people, but about getting people to behave in ways that further the terrorists’ goals.

It is, in part, this instrumentality of terrorism that distinguishes it from other forms of political violence. Take the Rwandan genocide in 1994, in which hundreds of thousands of Tutsis were slaughtered by Hutus seeking to ethnically cleanse Rwanda. Such slaughter is obviously political violence in the extreme. But the aim of the killings was not to spread fear amongst, or manipulate, the Tutsi population. Rather, the aim was to achieve the goal of ‘ethnic purity’ by systematically eradicating Tutsis. There was nothing symbolic about these killings – they were ends in themselves. This explains why we do not typically describe such killings as terrorism, even though they are clear examples of politically motivated violence.

So, that an action furthers a political agenda is not sufficient to make it an act of terrorism. But is a political agenda nonetheless a necessary condition of terrorism? Well, while an agenda of some sort seems necessary to capture the instrumental nature of terrorist attacks, we might question whether it needs to be a political agenda. Imagine that Jack the Ripper had the agenda of deterring prostitution in London. He might have hoped that by killing some prostitutes he could induce others to change their behaviour. If so, he certainly seems to have had an agenda. Does this mean that
we ought to classify him as a terrorist?

One argument against doing so is that there is a difference between trying to exert illicit influence on a political system and trying to influence the non-political behaviour of a small group of individuals. To reiterate an earlier point, many people take sovereignty – including political integrity – to be of immense importance. While what the Ripper did was very bad indeed, he did not try to manipulate the political integrity of a state for his own ends. His actions, even if successful, would have had much more limited ramifications than those of someone with a specifically political agenda. Of course, it doesn’t follow from this that his actions were morally any better or worse than those of a politically motivated terrorist. But they might nonetheless be different in kind to the actions of the politically motivated terrorist, because the Ripper’s actions do not undermine, or threaten to undermine, the sovereignty of the state in which he operates.

Of course, because we can draw this distinction between political and personal agendas, it does not follow that we ought to draw it. Whether we ought to draw the distinction will depend on how broad an account of terrorism we want to endorse. An advantage of a broad account will be that it is more likely to capture all of the actions that strike us as terrorism. But a disadvantage of such an account is that it will not capture only those things that strike us as terrorism. Counting as terrorism any violent action that furthers any agenda will lessen the moral force that attaches to the label ‘terrorism’.

Compare our use of the concept ‘torture’. If we define torture as simply ‘the intentional infliction of pain’, we are committed to a very broad account of torture that will probably include all those actions that strike us as torturous. On this account the abuse of detainees in Abu Ghraib will count as torture. But it will also include actions that don’t strike us as torturous, such as a mother smacking her child or a person inflicting defensive harm. Even if we think that smacking children is wrong, we probably don’t think that it is an act of torture. And since these intentional inflictions of pain are much less obviously wrong than the abuse of prisoners in Abu Ghraib, counting them as torture dilutes the moral connotations of the term itself. Saying that an action is torture will no longer tell us much about its moral content.

By opting for a narrower account of torture, we need not deny that non-torturous inflictions of pain can be wrong. Perhaps smacking one’s child is wrong. But the narrower account better captures the idea that torture is a distinctive sort of infliction of pain. Similarly, by opting for a narrower account of terrorism, which holds that only politically motivated actions count as terrorism, we can preserve its alleged distinctiveness.

Of course, not all politically motivated groups are terrorist. Indeed, we might think that not all politically motivated violent groups are terrorist. Armies are
politically motivated violent groups. It is precisely the desire to demarcate terrorism within the broader field of political violence that gives rise to attempts to identify something distinctive about the sort of political violence that terrorists perpetrate. With this in mind, let’s examine the other features listed above.

**Attacking non-combatants**

When we hear about terrorist attacks in the news, what is typically emphasised is that these were attacks that killed (or were intended to kill) 'innocent people.' In the context of war, the term 'innocent person' is often taken to be synonymous with 'non-combatant.' It is the fact that terrorists intentionally kill non-combatants that is generally thought to mark out terrorist killings from the collateral killings perpetrated by combatants in war.

As we saw in the previous chapter, it is not obvious that non-combatants can lay claim to innocence when their countries are engaged in fighting a war. Nor is it obvious that the distinction between intending/foreseeing makes a difference to an action's permissibility. These debates might influence our moral assessment of terrorist attacks. But that terrorists usually aim at civilian targets, rather than military targets, might still be a distinctive conceptual feature of terrorism. It seems to distinguish terrorism from combatant killings and from guerrilla warfare. But it does so by apparently conceding an important claim to terrorists, namely that they are engaged in genuine warfare. 'Non-combatant' is a military term (as is 'civilian'), and makes sense only when contrasted with combatants (who are, by implication, legitimate targets).

If we want to deny, as many people do, that terrorists are engaged in warfare, it doesn’t make much sense to offer an account of terrorism that relies upon these military distinctions. But if we remove these military terms from our analysis, we seem to be left with the claim that what is distinctive about terrorism is that it involves intentionally killing people. And this raises similar concerns to those noted above with respect to the broadness of our account. After all, ordinary murderers intentionally kill people all the time, as do state executioners and people engaged in self-defence. We do not want our account of terrorism to label all such killings terrorist.

Focusing exclusively on the targets of terrorist killings in order to identify terrorism’s distinctiveness thus seems to undermine the idea that there is something conceptually distinctive about terrorism, since the killing of innocent people is a feature of more than just terrorism. But if we combine the targeting of non-combatants with the criterion of political motivation, we may be better able to capture what is distinctive about terrorism. Just as euthanasia is distinguished from other kinds of intentional killing because of the agent’s benevolent motivation, perhaps terrorism is distinct from other kinds of killing (and other kinds of violence) because of both its distinctively political basis and its targeting of non-combatants.

However, it is worth noticing that there have been attacks upon combatants
that some people think qualify as terrorist attacks. For example, the group behind a planned attack on an Australian army base in 2009 were described as terrorists, even though it seemed clear that their plan involved killing only combatants. President Bill Clinton described the attack on the US Navy ship USS Cole in 2000 as an act of terrorism. Why would we call these attacks on obviously military targets acts of terrorism? The most obvious answer is that the attacks were sponsored and organised by terrorist groups – that is, groups that are responsible for other politically motivated attacks on non-military targets. Al-Qaeda claimed responsibility for the Cole attack, and the planned attack on the Australian base is alleged to have been backed by the Somali group al-Shabaah, who also have links to Al-Qaeda. Describing these attacks as terrorist implies that any attack carried out by a terrorist group – whether aimed at non-combatants or not – is itself terrorist. It also reflects the assumption that terrorism is about what we might call the pedigree of the perpetrators. The attack on the Cole was an act of terrorism not because of the nature of the target, but because of the nature of the attackers. It is to the idea of terrorism as non-state violence that we now turn.

**Non-state violence**

As we saw in Chapter 3, one of the *ad bellum* requirements of a just war is that the war be fought by a legitimate authority. We also noted that, for many people, this is a condition not only of a war being a just war, but of it being a war at all. Only legitimate authorities – usually understood as heads of state – have the capacity to declare war. If terrorists are not legitimate authorities, this might give us a way to clearly distinguish the killings carried out by terrorists from the killings carried out by regular combatants in war. Unlike regular combatants, who are employed by their state to carry out military operations, terrorists act independently: they are non-state actors. Perhaps terrorism can thus be characterised as unauthorised political violence.

The importance of legitimate authority stems from the idea that the state should have a monopoly on political violence. On entering society, citizens give up what Thomas Hobbes calls their ‘natural right’ to use violence, transferring this right to the state in return for the state’s provision of security and order. Thus, while states (and their representative agents) can legitimately use force, private individuals cannot. Of course, a private individual can use self-defence if necessary. But the explanation of the permissibility of self-defence is that, in a situation where my life is threatened and defence is necessary, the state has failed in its part of our bargain. It has not provided me with security, and thus I re-acquire my natural right to use force. With the exception of these cases, however, it is only the state that has a right to use force.

There is an obvious worry with defining terrorism as political violence carried out by non-state actors. As Robert E. Goodin puts it, such a definition ‘conceptually immunizes’ agents of the state against allegations of terrorism, making the idea that
state actors might engage in terrorism conceptually incoherent. Goodin points out that, for example, the definition of terrorism endorsed by the US government in the United States Code stipulates that "terrorism" means premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents" (emphasis added). This account of terrorism precludes, in principle, that a person acting on behalf of a national body – the state – might be guilty of terrorism:

The US Code's definition of 'terrorism' would have us think [that] states and their official agents can commit many wrongs strictly analogous to those committed by terrorists, perhaps; but they cannot commit the particular wrong peculiar to 'terrorism', which is (as with 'piracy') doing those things 'privately' rather than on the authority of some state.


Goodin's interpretation of the Americans' stance does seem to fit well with their frequent focus on the illegitimacy of terrorist actions (that terrorists are 'unlawful combatants'), which implies that the pedigree of terrorists is central to the concept of terrorism. Terrorists, unlike states, lack the proper authority to engage in political violence.

We can of course see why it would be advantageous for a state to endorse definitions of crimes that have built into them the notion that they are things that other people do. But using such a definition of terrorism is implausible for a number of reasons. For a start, it seems to involve mistaken assumptions about the scope of state authority. Notice that the US Code's definition of terrorism talks about violence aimed at non-combatant targets (a stipulation that is, incidentally, hard to reconcile with Clinton's assertion that the attack on the USS Cole was an act of terrorism). But violence aimed at non-combatants is prohibited by international law. If international law 'trumps' domestic law, it is hard to see how a state could properly be said to authorise such violence. States do not have the authority to override the laws of armed conflict. So even if a state actor is ordered to carry out such an attack by their superiors, this order cannot grant them the authority to carry it out, any more than you or could grant them that authority.

We might also object that it is simply false that states cannot engage in terrorism. Indeed, we might think that the vast majority of terrorism is perpetrated by states. States can engage in terrorism by sponsoring terrorist groups: by giving them financial support, trading arms with them, providing them with a safe haven for evading capture, and allowing the terrorists to set up training camps within their territory. Several countries, including the USA, have refused to recognise a distinction between terrorists and those who 'harbour' them in this way. If supporting a terrorist
group is enough to make one a terrorist, it seems that plenty of states engage in terrorism.

State-sponsored terrorism can also involve much more direct interaction with terrorists. Many people thought that the Libyan government orchestrated the 1998 Lockerbie bombing – the Libyan government itself eventually acknowledged responsibility and paid compensation to the victims’ relatives. But the involvement of the Libyan government – of the state actors – doesn’t seem to make the Lockerbie bombing any less an act of terrorism. Indeed, the USA has a ‘terrorism exception’ in its Foreign Sovereign Immunities Act (FSIA). The FSIA emerged out of trade disputes between US companies and overseas contractors. Under customary law, states were immune from prosecution in foreign courts. But as states became increasingly involved in international trade, their assumed independence was seen as giving states an unfair advantage in trade disputes. Companies could be held to legal account should they breach an agreement, but states could not. The USA, along with the UK, began to revise the law on prosecuting states, rejecting absolute immunity in favour of a more limited immunity. 5 And in 1996 Congress passed legislation permitting people to bring lawsuits in the USA against states that were responsible for terrorist acts involving US citizens, even if those attacks had occurred abroad. The very existence of this exception suggests that the USA is not consistently committed to the idea that only actions perpetrated by subnational or clandestine agents can count as terrorism.

The second way in which a state might engage in terrorism is by employing so-called state terror. Many states use violent secret police forces to routinely imprison, torture and kill their own citizens. These tactics are intended to spread fear amongst the population, and are usually politically motivated. Such tactics seem clearly terrorist in nature. However, as Samuel Scheffler points out, despite these similarities, state terror is different in character to ‘ordinary’ terrorism. Ordinary terrorism aims at destabilising a regime in order to effect change. State terror is almost always about bolstering control in order to maintain the status quo. Of course, preserving an existing regime is a manifestly political agenda, and presumably state terror has the same effect as ordinary terrorism, namely getting people to behave in ways that further the authors’ ends. But it isn’t clear whether state terror undermines sovereignty in the way that other forms of terrorism do. On the face of it, maintaining control seems like a way of preserving sovereignty, since sovereignty is all about power over one’s territory and political system.

However, perhaps state terror undermines sovereignty in a different way. Recall the idea of conditional sovereignty from our discussion of humanitarian intervention in Chapter 4. This was the view that states can jeopardise their claim to sovereignty by systemically violating or failing to protect their citizens’ human rights. This erosion of sovereignty paves the way for wars of intervention. States that engage in state terror
might undermine their sovereignty in this way – not by causing internal instability, but by lessening the strength of their claim that other states should not interfere in their affairs. If so, state terror can also threaten a state's sovereignty, albeit in a way unintended by the authors of the terror.

Thinking about state-sponsored terrorism and state terror makes it seem rather implausible that only those actions that are perpetrated by non-state actors can count as terrorism. While it may be true that many instances of terrorism are so perpetrated, it seems unlikely that this is what makes them terrorist. To borrow from Goodin, there is good reason to think that states can be terrorists, too.

**More violence**

It might seem obvious that terrorism must be violent. This doesn't mean that it must involve violence against people. But if we include blowing up buildings, vehicles and the like as violent acts, this will probably cover most instances of terrorism that spring to mind. We might also count certain attempts at violence as terrorist. Richard Reid's attempt to blow up a plane using explosives hidden in his shoes seems like terrorism, even though it was not, in the end, violent. This suggests that terrorism is a more 'holistic' concept than, say, murder. If a person tries and fails to commit murder, we say that they are guilty of attempted murder. But Reid is universally described not as an attempted terrorist but as a terrorist. When we identify something as terrorism, then, we seem to place a significant amount of weight on actions leading up to the final attack, classifying these preceding actions as terrorism in and of themselves. Because of this, however, it looks as if an action does not have to be violent in order to count as terrorism. Actions leading up to a terrorist attack – buying explosives, boarding a plane and so on – are not themselves violent, but most people nevertheless count them as terrorism.

Goodin agrees that not all terrorist actions must be violent. For example, warnings of terrorism can count as terrorism. Imagine that someone makes a phone call to a shopping centre, warning of a bomb that is about to go off. If the caller is a police officer intending to get the building evacuated in order to save lives, this hardly seems like an act of terrorism. But if the caller is issuing a bomb threat in order to cause fear and panic, this does seem like an act of terrorism, even though phone calls are not themselves violent. And making the phone call still seems like terrorism even if it turns out to be a hoax.

Goodin argues that this is because it is the agent's aims that distinguish acts of terrorism: 'What makes those warnings wrongful acts of terrorism is precisely the intention lying behind the acts, the intention to instil fear for political advantage.' His claim that making threatening phone calls about bombs in order to induce fear and panic counts as terrorism seems plausible (even in the absence of actual bombs). But Goodin argues that the range of non-violent acts of terrorism goes far beyond these
sorts of cases. He suggests that a significant amount of mainstream political rhetoric and argument counts as terrorism. There is a difference between warnings about terrorism, such as the call from the police officer just described, and terrorist warnings. Warnings about terrorism are intended to inform people of a threat in order to avert harm. Such warnings may generate fear, but that is not the intention behind their issue. Terroristic warnings, by contrast, aim at creating fear by reference to impending violence. And Goodin argues that some politicians issue warnings that aim to do just this.

This is not a new idea. Politicians are often accused by their opponents of ‘peddling the politics of fear’ – of trying to manipulate citizens into voting or behaving in particular ways by appealing to their anxieties. For example, the British government has been accused of playing on people’s fear of terrorism in order get Parliament to pass legislation undermining civil liberties, presenting this legislation as ‘counter-terrorism’. The government is not violent in its methods of manipulation, but Goodin thinks that such manipulation is nonetheless terrorism.

Goodin supports his claim by pointing out that while political leaders often insist that only violent attempts at political manipulation count as terrorism, other, more neutral, accounts do not endorse this view. For example, Goodin cites the Oxford English Dictionary’s definition of a terrorist:

‘One who entertains, professes or tries to awaken or spread a feeling of terror or alarm; an alarmist, a scaremonger’. So insofar as politicians’ warnings count as ‘alarmist’, they would indeed meet the dictionary’s definition of ‘terrorist’ ... The point is just this: people who do nothing more than issue warnings with the intention of instilling fear are in those other connections [e.g. the shopping centre case] deemed ‘terrorists’, provided they issue those warnings purely or predominantly for political purposes. Why not say the same of the public officials who do just the same?

(Goodin (2006), What’s Wrong with Terrorism?, pp. 99–100)

Goodin is not denying that violence is part of terrorism. What he is denying is that all acts of terrorism must themselves be violent. There are terrorist actions that are not themselves violent, but rather ‘piggyback’ on the violent actions of others.

So the hoax caller who issues the fake bomb threat at the shopping centre is engaged in terrorism, even though the phone call is non-violent, because the fear that the caller intends to create invokes or relies upon violence, albeit the violence of others. It is because terrorist groups have in the past planted bombs in shopping centres that bomb threats are effective tools of manipulation even when they are hoaxes. And, according to Goodin, politicians similarly rely upon the violence of others in creating fear that they can use for their own ends. It is because people are scared of
attacks by Al-Qaeda that they accept the erosion of certain civil liberties by their government.

Of course, this need not commit us to the claim that there are no moral differences between politicians who deliberately engender and manipulate fear and terrorists who go around blowing up buildings. We can distinguish, morally, between different sorts of terrorism. Terrorism that involves killing people is worse than terrorism that does not. But if what is crucial to terrorism is the use of fear as a political weapon, and if acts of terrorism need not be violent, there seems to be at least a prima facie case for thinking that Goodin is right that politicians often employ terrorist tactics to achieve their goals.

This broad claim about what counts as terrorism might raise concerns similar to those I noted earlier about diluting the content of terrorism. But, unlike the killings of Jack the Ripper, the kinds of actions that Goodin has in mind do seem to commit the particular wrong of undermining a state's political integrity. Thus we might think that while the difference between paradigmatic terrorists and the Ripper is a difference in kind, the difference between paradigmatic terrorists and terrorising politicians is one of degree.

Fear

Scheffler suggests that "if we define [terrorism] in a way that effaces or even breaks the connection between terrorism and terror ... then we are liable to miss some of the moral saliences toward which the word "terrorism" gestures. Surprisingly, it is the 'terror' part of terrorism that is often neglected in accounts of terrorism, which tend to focus more on the targets and political objectives of terrorists. But, as we have seen in Goodin's analysis, the intentional creation of fear as a weapon does seem to make terrorism conceptually distinctive. Other crimes may cause fear, but they usually do so incidentally – as a side-effect of pursuing some other goal. For terrorists the fear is crucial to their wider plans. The point of terrorism is to make people feel unsafe – to make them feel as if their government cannot protect them. This, in turn, can cause them to distrust their government and put pressure on it to bow to the terrorists' demands, or even seek to replace the government altogether.

Thinking about the sorts of targets that terrorists choose to attack, such as aeroplanes or high-profile political gatherings, can highlight the way in which terrorists deliberately cultivate fear. We all know, for example, that airport security is now incredibly rigorous. Every passenger is searched, scanned and questioned about their luggage. There are surveillance cameras, sniffer dogs and specialist police officers. Getting a bomb or other dangerous device onboard a plane is really quite difficult. And yet, even now, terrorists persist in their attempts to blow up planes. Why? Why target a plane, with all the inevitable obstacles, instead of a hospital or a cinema that will have
virtually no security and will be a comparatively easy target?

Disruption is one answer. Grounding planes causes massive inconvenience and costs governments and businesses money. But another answer, I suspect, is that hospitals and cinemas are too easy. It doesn’t take a great deal of skill, resources or knowledge to plant a bomb in a hospital. Getting a bomb onto a plane, in contrast, requires resources, expertise, contacts and training. By blowing up planes, terrorists let us know that even the very best security that our government can provide cannot stop them. No matter how sophisticated or stringent the security measures, terrorists can and do overcome them. And that is pretty frightening in its own right. It tells us that the terrorists are well-organised, clever and well-connected: they are part of larger organisations that train and fund them. They can blend in, undetected by the very people who are supposed to be watching for them and are specially trained to identify them. This assertion of the terrorists’ power is as much a part of generating fear as the bombings themselves, and it is rendered all the more apparent if someone manages to get a bomb onto a plane rather than attack an unprotected hospital.

Of course, the seemingly most obvious way in which terrorists create fear is by randomly killing innocent people. The sense of vulnerability created by random attacks is powerful indeed. But, in fact, terrorist killings are not often entirely random. In terms of the group that is targeted, terrorists are usually very specific. They have to be, if they want to further a specific political agenda. They will target the British, the Israelis, people working for particular companies or in particular industries, and so on, picking the group best placed to bring political pressure to bear in a way that will further the group’s ends. Where the process becomes random is with respect to the individuals within that group who are killed. By randomly killing within particular groups, terrorists make everyone who is a part of the group feel threatened, maximising the number of people who identify with the victims.

This technique is especially effective at creating fear when the targeted group is one that cannot be easily exited, or cannot be exited at all, such as a national or ethnic group. Since the British cannot help but be British, terrorist attacks that kill randomly within this group create a feeling of helplessness and persistent vulnerability amongst its members. It is these sorts of attacks that best fit Michael Walzer’s description of terrorism as attacks aimed at people not because of what they are doing, but because of who they are.

For example, attacks that target military bases or politicians do not tend to spread fear amongst the general population. They might give rise to a sense of indignation or anger. But most people do not feel as if they themselves are in danger, since they do not identify with the group perceived to be at risk. By contrast, when terrorists kill randomly within a group that cuts across social boundaries, perhaps by blowing up a train on the London Underground, they are able to cause significant
unrest amongst the society at large. As Scheffler says, ‘Figuratively and often literally, terrorism offers the biggest bang for one’s buck.’ Hitting a target that resonates across a significant cross-section of society enables one to maximise the impact of a single attack.

Is the intention to create fear a necessary condition of something’s counting as terrorism? I think Scheffler is probably right that accounts that overlook this feature are in some way bereft. The fear-based element is intrinsic to the workings of terrorism. Terrorists don’t achieve political goals by killing off anyone and everyone who stands in the way of those goals (as in the case of the Rwandan genocide). The killings are not eliminative in this way. Rather, terrorists kill some people in order to influence others, and this influence comes about primarily because people are afraid of what else the terrorists will do.

Imagine that an animal rights group kills a couple of politicians who were planning to vote against a ban on animal experimentation. The killings are intended to scare other politicians into voting in favour of the ban. This seems like terrorism: the killings are politically motivated and are intended to create fear as a means of manipulating behaviour. If, by contrast, the group simply murdered every member of Parliament who had planned to vote against the ban, this would be a very different method of achieving their goals. These people would be violent political activists. But I do not think they would be terrorists, because causing terror is not part of their plan. Such an approach would have more in common with the eliminative technique of genocide than the instrumental technique of terrorism.

WHAT’S WRONG WITH TERRORISM?

In the above discussion we examined various features of terrorist actions without taking a stance on the overall moral character of terrorism. Analysing the morality of terrorism is by no means an easy task. There are two main issues that such analyses usually seek to address. The first is whether terrorism is morally distinctive. To say that something is morally distinctive is not to say that it is worse than other actions, or that it is always wrong. Euthanasia is a morally distinctive form of killing – it is not like abortion and it is not like self-defence – but acknowledging this fact does not entail any claim about whether euthanasia is right or wrong, or whether it is better or worse than other forms of killing.

Of course, most of us probably think that terrorism is at least presumptively wrong. Terrorism often involves killing or maiming people who have not consented to be killed or maimed. Actions that cause these sorts of harm require strong moral justification. So the second issue that moral analyses of terrorism address is whether such justification is possible.

The notion of justified terrorism is, on some views, an oxymoron, akin to justified murder. Since the word murder simply means ‘wrongful killing’, it makes no
sense to wonder if murder might sometimes be right, because this amounts to asking whether something that is wrong might be right. Similarly, some people defend a view of terrorism in which the word 'terrorism' itself includes being unjustified as part of its meaning. We shall not explore these arguments here, but will rather keep an open mind about whether terrorism could, under certain circumstances, be justified.

**Killing non-combatants**

That terrorists intentionally kill non-combatants strikes most people as not only conceptually distinctive, but also morally distinctive. But, as mentioned above, it is the intentional aspect of the killings, rather than the killings themselves, that is the focus of most people's condemnation of terrorism. Terrorist killings are not like combatant killings, we think, because terrorists aim at non-combatants.

However, we have already seen that some people reject the moral significance of the distinction between intended harm and merely foreseen harm. Combatants know that they are going to kill non-combatants. This is why collateral killings are described as incidental, rather than accidental. Of course, they do not intend the killings, and intentions often play a role in our moral judgements of actions. But, as Lionel McPherson argues, it is not obvious that intentions are so important that they can sustain the alleged moral distinction between terrorism and collateral damage:

Acts of conventional war ... are not as susceptible to evaluation through this feature of commonsense morality. The harm done to noncombatants through many of these acts is likely, foreseeable, avoidable, and extensive – which would appear to largely overshadow the relevance of the combatant's intentions to permissibility. Commonsense morality recognizes that agents might not be morally culpable when, despite what they reasonably could expect, they do unwarranted harm. But when the unwarranted harm can reasonably be expected, commonsense morality is not committed to recognizing that the agents' intentions make a moral difference...


McPherson suggests that even if we allow that intention is sometimes morally relevant, it does not follow that it always makes the difference between permissibility and impermissibility.

This becomes especially apparent when we think of those who generate unintended harm as a side-effect of pursuing a morally bad or neutral goal. Consider a factory owner who dumps his factory's waste into a local river because this is cheaper than disposing of it safely. His intention is only to save money. He does not intend to pollute the river and cause environmental damage. But few of us would think that this
matters very much for our moral judgement of his action. Provided that he knew that
the pollution was likely or even possible, we would probably hold him just as
responsible for the damage as if he had intended it. That polluting the river was a
side-effect of his action, rather than his end, does not make polluting the river
permissible.

If so, it seems that there are times when intention makes little or no difference
to the morality – and permissibility – of an action. As we saw in Chapter 7, some just
war theorists argue that causing any foreseeable harm to non-combatants is
impermissible – that the distinction between intending and foreseeing is simply
irrelevant. McPherson’s own view is that the difference in intentions simply pales into
insignificance alongside the magnitude of the harms being inflicted. If he is right,
invoking the difference in intentions between terrorists and combatants is not enough
to show that terrorist killings are wrong in a way that collateral killings are not.

**Legitimate authority and representative authority**

However, McPherson is not claiming that terrorism is therefore permissible, or that all
collateral killings are therefore impermissible. McPherson offers another explanation
for why terrorism is wrong in a way that the violence of regular combatants is not. We
have already encountered the objection that terrorists lack the legitimate authority to
employ political violence. But McPherson argues that this objection does not quite hit
the nail on the head. There is, he suggests, an important difference between a
legitimate authority and a representative authority. The criteria for legitimate authority
make a poor vantage point from which to criticise terrorism because these criteria
themselves lack moral credibility. One can, for example, be a legitimate authority by the
lights of a Hobbesian social contract theory simply because one provides order and
security for one’s citizens. But this very minimal condition for legitimate authority
confers legitimacy on even the most morally bankrupt of states. A harsh totalitarian
regime can offer security for its people and count as a legitimate authority. But can we
really judge the rightness or wrongness of terrorism by whether it fits with a notion of
legitimacy that confers authority upon those sorts of states? McPherson thinks not.
Failing to satisfy the ‘morally weak, Hobbesian account of legitimate authority hardly
seems a compelling reason for judging that nonstate terrorism is wrong’.

What matters for the legitimacy of political violence is not whether one is a
legitimate authority in the Hobbesian sense. Rather, McPherson argues, what matters is
whether one can lay claim to representative authority, which results from an
appropriate process of consultation with the people that one claims to represent:

The deeply distinctive problem for nonstate terrorists now emerges. That they
lack legitimate authority is only a rough indication of the problem. Political
violence by nonstate actors is objectionable when they employ it on their own
initiative, so that their political goals, their violent methods, and, ultimately, their claim to rightful use of force do not go through any process of relevant public review and endorsement. Nonstate terrorism's distinctive wrongness does not lie in the terrorism but rather in the resort to political violence without adequate licence from a people on whose behalf the violence is purportedly undertaken.

(McPherson (2007), 'Is terrorism distinctively wrong?', p. 542)

Representative authority matters, in McPherson's view, because peoples are entitled to self-determination. Even in less than ideal states, there are usually processes of consultation that allow the citizens to influence how their state behaves. Thus, in the case of regular combatants, we can usually assume that they carry out the will of at least a majority of those whom they represent. But terrorist groups who claim to represent a people rarely give the members of that people the opportunity to endorse or reject them. McPherson's argument against terrorism is thus broadly Kantian, stressing that respect for peoples' autonomy requires allowing them to control what is done in their name. Terrorists who employ violence 'on behalf of' a people without properly consulting that people fail to respect its members' autonomous rights.

McPherson is not arguing that terrorist groups can never have representative authority of the sort that he describes. Some terrorists groups, such as the African National Congress (ANC), could lay claim to such authority. But that one is a representative authority is not sufficient to make the use of political violence permissible. No group – state or otherwise – is entitled to use force simply because it represents a people. Most terrorist groups not only lack representative authority, but also lack a just cause. However, like Scheffler, McPherson does not rule out the possibility that a terrorist group might have such a cause and that, under certain circumstances, their use of terrorism might be permissible.

One difficulty with McPherson's account is that of discerning which groups count as 'peoples' of the sort that can confer authority upon their members to violently pursue their goals. McPherson grants that such peoples can emerge as a result of 'self-ascribed nationality, ethnicity, culture, or religion, or on the basis of being victims of common oppressors'. But if we allow a great many shared features to form the basis of a people, nothing prevents fundamentalist Islamic militants, for example, from claiming this status. After all, they are all fundamentalist Islamic militants – they share this particular cultural or religious feature. It looks as if, by McPherson's lights, such a group can qualify as a people and bestow representative authority upon themselves.

While acknowledging this worry, McPherson suggests that it is not as serious as it first appears. Terrorist groups rarely claim to represent only a small minority, but tend to rather grander assertions about representing, for example, all Muslims or all Palestinians. McPherson argues that this is precisely because such claims give a more
plausible impression of ‘representative authority that is morally compelling’. If this is the case, it implies that terrorist groups themselves endorse the idea that they must have wider endorsement than merely their own ‘parochial source that answers only to the edicts of leaders who lack relevant public approval’.

But it’s not clear that this will satisfy those who level this objection. For a start, it seems a merely contingent fact that terrorists often make these claims. Such a fact cannot rebut the theoretical objection that McPherson’s account gives us no way in which to restrict what counts as a people. Moreover, there are explanations of terrorist proclamations of representation that do not support McPherson’s analysis. It seems possible, for example, that those terrorists who claim to represent all Muslims might really believe that they represent these people, and that their methods are approved by them. But this doesn’t mean that they take their actions to be morally permissible because of this approval. It seems more likely that they think them permissible because of the ‘parochial source’ to which McPherson refers. This is the source of the morally credible licence. They just happen to believe (mistakenly) that other Muslims endorse this source too. So we cannot be so quick to assume that terrorists accept the need for endorsement by those whom they claim to represent. They may simply assume this endorsement because of prior assumptions that the Qur’an confers legitimacy on their actions, and that all Muslims endorse the Qur’an. For these people, there may not be any relevant public approval that invoking the Qur’an fails to satisfy.

Furthermore, only some terrorist groups seek to present themselves as acting on behalf of a wider people. Others, like animal rights activists and environmental terrorists, do not claim to act on behalf of people at all. Who ought to be consulted about the decision to use violence to further the cause of animal welfare? Who should have the power to sanction or prohibit violence in the name of this cause? In so far as we can make reasonable judgements about what the animals themselves would prefer, the onus would seem to be on those opposing the activists to show that their actions are not genuinely representative of the animals’ preferences.

There are still other terrorists who do not claim to represent a people because they take their orders directly from some divine being (thinking that they are God’s messengers on earth, for example). If we assume, for the sake of argument, that there is no God, such a group or person cannot commit the wrong of violating God’s autonomy by failing to engage him or her in a proper process of consultation. If what is distinctively wrong with terrorism is that terrorists claim to represent a group while failing to appropriately consult that group, thereby violating that group’s rights of autonomy and self-determination, there will be a significant number of terrorist groups who do nothing distinctively wrong (although, should they lack a just cause, their action will still be wrong in the way that other unjust violence is wrong). Only those who lay claim to wide representation while failing to genuinely engage with that group will
undermine that group's autonomous rights in the way that McPherson describes.

**Using as a mere means**

McPherson’s rejection of intention as a way of distinguishing between terrorism and collateral damage also warrants closer attention. For example, Scheffler argues that it is not simply that terrorist killings are intended that makes them morally different from combatant killings. The distinctive wrong of terrorism is more specific. Terrorists don’t just intentionally kill noncombatants – they also make use of their victims in a morally distinctive way.

Recall the earlier description of how terrorists typically choose their targets – that they select a specific group and then randomly kill a subset of that group. Scheffler describes the members of this subset as the terrorists’ primary victims. But these victims are not only intentionally killed. They are also used: it is their deaths that enable the terrorists to spread fear and instability amongst the remaining members of the target group. Scheffler calls these remaining members of the target group the secondary victims of terrorism, because it is they whom the terrorists intend to terrorise. These secondary victims are also used as mere means in their turn, since it is they who are manipulated into bringing political pressure to bear on their leaders. The secondary victims, just like the primary victims, are crucial tools in achieving the terrorists’ goals. Scheffler thus argues that terrorism is morally distinctive not simply because terrorists intend the deaths of their primary victims, but because they use those victims as a means to a means, an extreme violation of the Kantian prohibition on using people.

Kant argued that we should always treat people as ends in themselves, respecting them as autonomous individuals with their own projects and goals. If I go on a date with you just to make somebody else jealous, I fail to treat you as an end in this Kantian sense. By regarding you merely as a way of getting what I want, I fail to accord you the proper respect that you deserve as a human being. I don't think about your goals or projects, but only about how you can help me further mine. You could rightly complain that in dating you, I was just using you. Terrorism, according to Scheffler, abuses people in the extreme of this Kantian sense. All members of the target group are viewed as tools to be manipulated in the pursuit of some goal that they do not share. Such treatment is the epitome of lack of respect for persons:

Those who engage in this kind of terrorism do not merely display callous indifference to the grief, fear and misery of the secondary victims; instead, they deliberately use violence to cultivate and prey on these reactions. This helps to explain why there is something distinctively repellent about terrorism, both morally and humanely.

(S. Scheffler (2006), 'Is terrorism morally distinctive?', *Journal of Political Philosophy* 14, No. 1, 1-17, p.10)
By deliberately engendering 'fear and misery' terrorists do something not only morally distinctive, but, in Scheffler's eyes, morally abhorrent. When terrorists mount a sustained campaign of attacks, they create an environment where people live in perpetual fear, unable to carry out their normal lives. Scheffler echoes Hobbes' claim that 'Continual fear – not momentary anxiety but the grinding, unrelenting fear of imminent violent death – is unspeakably awful ... A life of continual fear is scarcely a life at all.' Scheffler does not, however, argue that terrorism is always wrong, even though he does think it a prima facie evil. He leaves it open whether there could be circumstances that are themselves so awful that the use of terrorism is justified as a way of overcoming those circumstances.

Goodin's account of the moral distinctiveness of terrorism is also broadly Kantian and reflects his claim that the conceptual distinctiveness of terrorism comes from the use of fear to achieve political goals. Goodin argues that fear-mongering in this way subverts citizens' capacity to reason clearly, causing them to vote or lobby reactively rather than rationally. In other words, terrorists erode what Kant identifies as most valuable about people, namely their 'capacity for autonomous self-government, both individually and collectively.'

Goodin describes this as a primarily political wrong rather than a moral wrong, although of course the two are not mutually exclusive. Terrorism is morally distinctive, and morally bad, because it threatens our ability to participate rationally in the state, thereby threatening our ability to shape our own lives. Of course, neither Scheffler nor Goodin deny that much of the moral gravity of terrorism comes from the fact that terrorists intentionally kill people. It is the killing that makes terrorism morally serious. But Kantian analyses can show us why terrorism also seems morally different.

**Permissible terrorism?**

Tony Coady argues that terrorism is clearly wrong. But, he says, it may be that terrorism is like lying – usually wrong but sometimes justified: 'The really interesting moral issue is thus not whether terrorism is generally wrong, but whether it is nonetheless sometimes morally permissible.' As we have seen, several contemporary just war theorists are at least open to the idea that terrorism could, under certain circumstances, be justified. After all, most people do accept that it can be permissible to incidentally kill some noncombatants in a just war. If non-combatant lives can be weighed against just ends, it seems possible that they can be so weighed not only by states, but also by terrorists. And even if we think that intentional killing requires much more justification than incidental killing, this is compatible with thinking such justification possible. The threshold, though high, might yet be attainable. So what sort of causes might be so morally important as to warrant intentional killing of the sort that terrorists employ?

Saul Smilansky argues that, in the absence of gratuitous poverty or severe oppression, merely being under foreign rule cannot justify terrorism. Smilansky's view
echoes the ideas put forth in Chapter 3 regarding proportionality and just cause. Recall May’s and McMahan’s claims that even if a nation suffers the appropriate sort of wrong to have a just cause for war, such as the illegal annexing of its land, war can nevertheless be disproportionate and therefore unjust. Similarly, even if the group whom the terrorists represent (and might genuinely represent in the way that McPherson requires) have suffered an appropriate wrong, it does not follow that terrorism is a proportionate response. Smilansky argues that to have a chance of being justified, terrorism must be employed for a proportionate just cause and be a last resort after all other means have been exhausted. As an indication of the sort of historical causes which might have met these conditions, Smilansky lists the ‘almost genocidal’ situations of Biafra, Cambodia, Rwanda, Sudan and East Timor. These are cases in which the cause is clearly proportionate, since ‘there is clear danger to a group’s very existence or [of] the mass extermination of noncombatants’.

Smilansky also suggests that terrorism might be warranted as a means of drawing the world’s attention to poverty in the Developing World, where the catastrophic effects of famine and deprivation have been allowed to rage largely unchecked. If a state inflicted this amount of suffering upon another state in an act of aggression, this would be a proportionate cause for war. And since the suffering is so very great, it could plausibly serve as a proportionate cause for terrorism as well.

Finally, Smilansky suggests that terrorism could be a permissible means of toppling tyrannical regimes. Where a regime is undemocratic and severely oppressive, it is often only by forcefully ousting its leaders that change can be brought about. In such cases Smilansky suggests that ‘limited and narrowly focused terrorism’ could be justified (assuming that such force had a good chance of being effective).

However, Smilansky concludes that what is interesting about terrorism is that it is widely employed in cases where it is obviously unjustified, and only rarely employed in cases where it would be justified. People have used terrorism extensively in the name of religious extremism, but not, for example, to try to galvanise the West into combating poverty in the Developing World:

Our result implies that the world is curiously disjointed. Perhaps there are situations where terrorism has been contemplated but not pursued as a result of good moral reasoning. Still, in a striking way the role of adequate moral reflection is shown in its emptiness – both when the efforts at justification ought to yield negative results and when they ought perhaps to yield positive ones. Within the societies and cultures that have generated terrorism, or support it, moral deliberation has failed to be effective. The thought that terrorism can be adequately guided by processes of justification is an illusion ...
hope of ‘fine tuning’ the permission of terrorist activity is far too risky. We should err on the side of not allowing terrorism.

(Smilansky (2004), 'Terrorism, justification and illusion', pp. 799-800)

Smilansky's claim is that we do grave harm, and no good, by entertaining the possibility that terrorism can be justified. By failing to unequivocally condemn terrorism, we give succour to those who 'justify' their causes by some parody of moral deliberation. Smilansky suggests that the sort of people who resort to terrorism are, by and large, the sort of people who bypass moral reasoning altogether, being motivated instead by religious fanaticism and ancient feuds. He argues that we are obliged, given the apparent irrelevance of justification in cases of actual terrorism, to denounce terrorism across the board, lest those with malevolent causes make use of our arguments for their own ends. He cites Bernard Williams's observation that engaging in moral philosophy is itself morally perilous, since 'we may run the risk of misleading people on important matters'.

While it may be true as a matter of empirical fact that terrorists abuse the notion of justification for their own ends, Smilansky's thoughts on how we should respond to this abuse are both surprising and, I think, unattractive. Of course, we can be mindful of the practical implications of advancing particular views. But there are a great many arguments in moral philosophy that are open to abuse. Offering justifications for euthanasia runs the risk that one's arguments will be gleefully adopted by some grasping relative keen to bump off granny before her time. But even if this might be a good argument for making euthanasia illegal, it is not a good argument for prohibiting moral discussion of euthanasia.

Moreover, if we're talking about practical implications, Smilansky's own view is that the sort of people who engage in terrorism don't really care about moral justification, being motivated instead by religious fanaticism and ancient feuds. But such people are then unlikely to be any more swayed by an unequivocal condemning of terrorism than they are by the revelation that their cause in fact fails the justification test.

The sort of people who might care about justification, however, are the sort of people whom Smilansky identifies as having a just cause, such as those facing genocide or horrific oppression. Such people are not motivated by religious fundamentalism or nationalistic hatred, but rather by a desperate need to defend their own lives. A case that is notable by its absence in Smilansky's discussion is that of the ANC, whose use of terrorism in a bid to overthrow the South African government and end apartheid seems to fit exactly Smilansky's description of limited and narrowly focused terrorism aimed at abolishing tyranny. And the ANC began as a non-violent group, becoming increasingly militant only after peaceful means failed to produce results.
Given the terrible conditions imposed upon black South Africans during apartheid, and given that terrorism did seem to be a last resort in this case, many people think that it is precisely cases like this that make it worth discussing whether terrorism can be justified. And given Smilansky’s own acknowledgement that terrorism is sometimes justified, and his claim that unjustified terrorists don’t care about morality anyway, it seems that, by his lights, we can only make things worse by universally condemning terrorism. Doing so will not discourage those who lack just causes, but it may discourage those whose cause is just.
CHAPTER 2

Duties of Humanity
On January 12, 2010, an earthquake hit the small Caribbean island of Hispaniola, home to the Dominican Republic and Haiti. Before the earthquake, Haiti was by almost all indicators one of the poorest countries in the world. From child and infant mortality to malnutrition, life expectancy, and income, Haiti was the worst-off country in the Western Hemisphere. At 4:53 p.m., the already precarious livelihoods of many Haitians were shattered by a 7.0 magnitude earthquake whose epicenter was just outside the capital of Port-au-Prince.

The immense suffering caused by the earthquake lasted long after the buildings collapsed. The first phase of suffering was a result of the direct devastation from the earthquake itself. Because most physical structures in Haiti were constructed out of rudimentary materials without the reinforcements that would be able to withstand a major earthquake, somewhere between 100,000 and 200,000 victims were killed in the earthquake. Without adequate infrastructure, or equipment to guide the rescue, people who were injured but not yet deceased had little chance of immediate rescue. In a second phase, which in many ways continues today, widespread suffering befell Haitians because of the after-effects of the earthquake—lacking adequate shelter, employment opportunities, a public health system, and public security, Haitians lived in insecure housing, faced bleak economic prospects, went hungry, and suffered from the spread of disease (including an outbreak of cholera brought by UN troops who were supposed to bring stability to the country) and were at risk of physical and sexual violence. The Haitian government was a paradigmatic extractive institution, lacking capacity to deliver public goods and systems of accountability to ensure genuine responsiveness to democratic demands in rebuilding.

As news of the devastation spread across the world, the call for help went out to individuals, governments, development agencies, and international institutions. Haitians were in dire need. Now was the time to lend a helping hand. Celebrities, politicians, and average citizens held benefit concerts, rallies, and undertook unprecedented fundraising efforts. Phone companies set up systems that allowed people to give quickly and directly to humanitarian relief organizations. The shared sentiment among the international community was clear—we all must chip in to help Haitians in need. In the months and years following the earthquake, upwards of 9 billion USD was raised to help Haiti ‘build back better’. The Red Cross alone received half a billion dollars in donations.

The Case for Duties of Humanity

What moral reasons might justify this call to action? In 1972, Peter Singer published one of the most famous articles in contemporary philosophy. In it, he considered catastrophe occurring at the time, a famine in what is today Bangladesh. In ‘Famine,
Affluence, and Morality,’ Singer provided the following thought experiment. Suppose that you are on your way to work, perhaps to teach a class or attend an important meeting or start the assembly line. On your way, you come across a child drowning in shallow pond. As an adult, you could easily enter the pond and save the child from drowning. You may incur minor costs for doing so: you may need to pay dry cleaning for the clothes you have soiled, or suffer minor repercussions at your workplace for arriving late. What should you do?

Singer, like most of us, believes that you are obviously morally required to save the drowning child. No minimally adequate morality would allow us to blithely oversee the death of an innocent child simply because we did not want to pay a small dry cleaning bill or be a bit late for work. Singer thinks a single moral principle explains our intuition in this case—namely, that if we can prevent something very bad at minimal cost, we are morally required to do it.

Relying on the intuition generated by the shallow pond case, Singer constructs a simple three-premise argument that, if correct, overturns widely shared beliefs about our duties to reduce global poverty.

First premise: Suffering and death from lack of food, shelter, and medical care are bad.

Second premise: If it is in your power to prevent something bad from happening, without sacrificing anything nearly as important, it is wrong not to do so.

Third premise: By donating to aid agencies, you can prevent suffering and death from lack of food, shelter, and medical care, without sacrificing anything nearly as important.

Conclusion: Therefore, if you do not donate to aid agencies, you are doing something wrong.

(Singer 2009, p. 15)

Singer’s argument, in its initial formulation, involves a strong (or very demanding) version and a weak (or less demanding) version. The strong version states that it is morally wrong to fail to save a life up until the point at which you are giving up something of morally comparable value to that which would be saved. To consider how stringent our duties of beneficence would be on this formulation, think about the things that may exist in your life that you could consider giving up in order to relieve global poverty. At what point would something that you have to give up be morally equivalent to saving a person’s life? Upon reflection, something of equivalent moral weight will be very significant indeed. In fact, it may require giving away all of your money up until the point at which you are at the point of minimal subsistence, perhaps just at the basic consumption level at which you can continue to earn income to give
away to charity.

A weaker version of Singer’s position would only require that you donate financial resources to aid agencies up until the point at which you would be giving away something of moral significance. In Singer’s most recent work on the topic, he seems to concede that most people will not give up most of their salary for anti-poverty work. So he recommends a sliding scale by which high-income earners can decide what a minimally adequate scheme of donations would require. Singer suggests that individuals earning over 100,000 USD give 5 percent of their income to charity, with incremental increases the higher the annual wages go, up to the point where individuals making over 10,000,000 USD give one-third of their annual income to charity.

Singer is a proponent of utilitarianism. Utilitarians are committed to the view that moral agents ought to select that action which will, in aggregate, maximize the good. The good is defined as individual utility, which for Singer is measured as either the degree to which someone satisfies their preferences or the degree to which they experience pleasure and avoid pain. Utilitarianism certainly supports the view that we ought to give much of our money to anti-poverty work. There is much greater marginal utility to be gained by the recipient of a given portion of foreign aid than there is marginal utility to be lost by the person who forgoes an expensive meal out or a fancy new shirt to provide that foreign aid. However, while Singer is a utilitarian, he believes that the argument he has provided does not rely on the endorsement of utilitarianism. The duties of humanity (also commonly known as duties of beneficence or duties of assistance) he proposes are meant to be ecumenical, in the sense that they could be (and on his view ought to be) endorsed by non-utilitarians.

Despite the initial simplicity and plausibility of Singer’s argument for humanitarian duties to alleviate global poverty, Singer’s argument faces a range of objections that threaten his conclusion that wealthy individuals have very demanding obligations to reduce global poverty.

**Competing Intuitions and Competing Principles**

Singer’s argument relies on the intuition that it is morally wrong to fail to save the drowning child. This intuition is quite strong in the case he presents us. But it conflicts with other intuitions we have. When asked, most people do not think they have an obligation, let alone one as stringent as Singer suggests, to give away all of their wealth and a significant portion of their income to poverty alleviation. (Even if they did think this, their behavior indicates otherwise, whereas, with other moral duties, such as the duty to not kill innocent people, most people both think this is a strict moral duty and act in a way that respects the duty.) Of course, Singer believes that the intuition in the case of the drowning child is the one we should trust … not our views on charitable giving. But why is that? Perhaps our selfishness, or our embrace of consumerism, or our
misguided desire to match the living standards of our peers have convinced us that most of the money we earn should remain ours. Yet if we are able to set aside these confounding factors, Singer thinks we should realize that we really do have strict duties of beneficence to prevent suffering around the world.

However, when considering cases that appear similar in all relevant respects to the drowning child, most people generate different intuitions about whether we have a duty to rescue those in need. Gerhard Øverland and Christian Barry ask us to consider a similar example:

Bob is sitting in his house doing some Internet banking. Unbeknownst to his neighbours (the Smiths) he can see and hear them through an open door in the veranda. He notices that they are discussing the state of their terminally sick child, Jimmy. They need new and expensive treatment to save Jimmy ... Clicking over the money would save Jimmy, but most of Bob’s savings for retirement would be gone. Bob decides not to click the mouse.

(Barry and Øverland 2009, p. 241)

In this case, most people do not think that Bob acts wrongly if he does not transfer over the money. And this case seems to have all of the same features as Singer’s drowning child. So then it appears that there is a clash of intuitions, each apparently widely held and reliably produced. Given these competing and strongly held intuitions, it becomes unclear whether we should endorse the general principle that ‘whenever one can prevent suffering from happening without giving up something of comparable moral cost, one should do so.’ If the principle cannot account for our conflicting intuitions in competing cases, we may have reason to weaken or modify the principle in a way that still explains our views in the drowning child case but also issues the correct judgments in other cases.

For example, a competing principle might state, ‘When one can prevent something bad from happening, one’s relationship to the bad activity is clear; others are not well position to provide assistance. If one is guaranteed success in preventing the harm, at minimal cost to oneself, one ought to do it.’ This alternative principle requires that a person be clearly related (such as being physically proximate or directly able to provide assistance) to the morally problematic situation, that they be assured of success in preventing the harm, that they be the only agent capable of providing assistance, and that the action be minimally costly to them. This principle still generates a requirement that the child be saved from drowning, but does not require extensive donations to relieve global poverty.

**Demandingness and Fair Shares**

A different way to reject Singer’s argument is to claim that his conclusion is overly
demanding. On this view, our moral principles must make only modest demands on us as moral agents. They cannot require that we be moral saints, or that we act beyond a standard of what a reasonable person could be expected to do. The demandingness objection is certainly what a lot of people have in mind when they try to reject Singer’s argument. Surely it cannot be wrong for me to own a car, or to pay to send my children to a good university, or to occasionally enjoy a nice dinner and a movie with my partner.

Yet it is not clear that over-demandingness is in itself a reason to reject a moral theory or moral principle. Morality requires what morality requires, and in some cases the requirement might be for very great sacrifices indeed. Some individuals, for example, might be morally obligated to give their lives to protect their children from an impending threat. The mere fact that this is an exceedingly demanding moral obligation does not prove that it is not a moral obligation. If it turns out that basic facts about the world or basic features of morality require a great deal of sacrifice from moral agents, then that is what morality requires.

One way to respond to apparently overly demanding accounts of morality is to argue that morality must take account of the sorts of agents that human beings are and could be. Even if morality sometimes demands great sacrifices, a general morality that constantly requires extremely demanding duties of beneficence is not one that can reasonably be adopted by normal human beings.

A second objection to Singer’s strict duties of beneficence holds that duties of beneficence are imperfect. Imperfect duties are those duties that we are required to discharge some of the time to some extent. But unlike strict duties, which we may never violate, imperfect duties may be overridden by even self-interested, non-moral reasons on at least some occasions. On this view, people are required to assist in the eradication of poverty, but only to a point. Proponents of this weaker view must then specify some threshold at which a moral agent has discharged her moral duties to eradicate poverty.

A third and related approach, argued most prominently by Liam Murphy, defends the idea that in circumstances in which not all individuals discharge their duties of assistance, individuals are only morally required to do as much as they would need to do if everyone did their share. If my neighborhood street is strewn with garbage after a festival, my fair share duty is to help to pick up that amount which, if we all do some garbage collecting, would leave the street clean. I am not morally required to clean the whole street.

An attraction of the fair share view is that it retains a sense of fairness about how much of a sacrifice people should make to solve a problem, given that other people may not comply with what morality requires of them. That is, I am obligated to do my part to solve the problem, but I don’t have to take up the slack indefinitely while others fail to discharge their moral obligations. This requires establishing the threshold
at which one has done their fair share. One way to identify this threshold is to figure out what it would take for the problem under consideration to be solved under conditions of full compliance. Suppose, for example, that you determine the amount of financial resources it would take to lift everyone out of global poverty. Then divide this figure by the number of people who are currently in a position to provide assistance. This is how much you are morally required to give.

This response is attractive because it retains a sense of humanitarian obligations to others while defusing worries about over-demandingness. But it presents new problems as well. Given that it is foreseeable that some people will not comply with their duties, why should the moral demands on complying agents be weakened, since doing so will result in avoidable suffering for the potential recipients of their efforts?

Consider the Global Fund to Fight AIDS, Tuberculosis, and Malaria. This fund currently receives donations from a number of countries. It is foreseeable that many other countries will not donate their fair share to the fund to combat these diseases. But morally reflective governments might take this foreseeable non-compliance by the uncharitable as a reason to increase their donations. When a collection of agents seeks to achieve some goal, and some of those agents don’t hold up their end of the bargain, the moral importance of achieving the goal is unchanged. It may be, therefore, that it is precisely in circumstances of non-compliance with moral duties that we are morally required to do more than our fair share.

**Idealization and Reductionism**

A different set of objections focus more squarely on Singer’s way of framing the problem of global poverty rather than focusing on his conclusions regarding duties of assistance. Singer’s argument makes idealizing assumptions about the nature of global poverty that problematically masks some of its most important features.

Prominent among these oversights is the mistaken view that poor people are simply moral patients rather than being moral agents. The child drowning in the pond is helpless, and presumably incapable of exercising agency. But poor people’s lives are nothing like this. They struggle daily to put food on the table and to make their lives better. Just like you and I, they make mistakes and poor decisions, but on average they are working hard, coming up with creative solutions to improve the lives of themselves and their families despite the bad misfortune of being born into environments in which material and social deprivation persist.

The history of international development work, as we will discuss in Part III, is littered with projects that have failed because they did not have the consent, or support, of their intended beneficiaries, or were insensitive to the unique social and political context in which deprivation persists. The architects of these failed projects
were very much in a Singer-style mindset: ‘These people are poor, this project will lift them out of poverty, we must save them!’ By failing to take seriously the agency of individuals living in poverty, and being insensitive to the underlying political mechanisms that produce deprivation, Singer mistakenly treats poverty reduction as a mere matter of wealthy saviors assisting people in poverty.

People are poor because they are deprived of the individual and collective power to secure their basic rights and unleash their creative potential. A diagnosis of the problem of poverty that is insensitive to the power dynamics and institutional arrangements that produce deprivation is likely to fail to produce an adequate solution. Moreover, it is morally objectionable to talk about and treat people living in poverty as if they are not moral agents in their own right, capable of adapting to and struggling against highly unjust circumstances. Such disrespect itself deserves moral condemnation, even if the disrespect is not intended and is part of broader efforts to provide assistance.

**Poverty Is No Pond**

Following the earthquake in Haiti with which we started this chapter, there were unanimous pledges to help the Haitian people rebuild. The motto was to build back better, with Haitians at the center of the rebuilding effort. Four years on (at the time of writing), we are able to assess how this aid effort was undertaken. Unfortunately, providing anti-poverty assistance in Haiti is much more difficult than saving a drowning child.

While some of the donations to rebuild Haiti have done some good, it is extremely difficult to say that the aid effort was successful in its pledge to ‘build back better’. The government of Haiti received less than 1 percent of immediate humanitarian relief and somewhere between 15 percent and 20 percent of longer term relief funds. While the government is extractive and corrupt, it is extremely difficult to see how institutional improvement can occur, and a government can function at all, if an aid program is run almost entirely outside of its control.

The Red Cross raised over a half a billion dollars to provide assistance in Haiti, more than any other charity. In one project investigated by ProPublica, they claimed to have built 130,000 homes in a neighbourhood outside Port-au-Prince, but the number of permanent homes that had been constructed was six (Elliot and Sullivan 2015). The Red Cross has saved millions lives throughout its history. Why did it struggle in Haiti? The reasons for their failures are familiar to critics of foreign assistance: they relied heavily on employees who did not speak the local language, understand the domestic context, or respect the intended beneficiaries of their projects. Constraints to development prior to the earthquake, such as an inadequate system protecting and regulating property rights in land, made it difficult to proceed with housing projects. And they lacked local partnerships and political buy-in that would allow for successful operations on the ground.
As Leif Wenar argues, poverty is no pond. International development efforts face a range of challenges that make it difficult to successfully promote poverty reduction. It is difficult to coordinate the many different actors, often with conflicting and overlapping agendas, involved in aid provision; it is difficult to enlist local participation in support of projects while often aiming to change local practices; it is difficult to prevent resources from being diverted; aid may have deleterious economic effects, including freeing up government resources for more harmful activities like arms purchases while also undermining local economies; aid projects may overlap, thereby undermining the effectiveness of each individual project; aid may weaken demand for accountable governance; and aid may wrongly impose certain external values to the detriment of locally valued customs.

Thus, it is no surprise that, according to one of the most careful analysts of relief efforts in Haiti, today,

Several thousand people still live in tents that were supposed to be temporary housing for families displaced in January 2010. Haiti remains the 'Republic of NGOs.' Five years and $9 billion later, Port-au-Prince does not have decent roads, clean water, or a reliable supply of electricity.

(Ramachandran 2015)

**A Better Way for Duties of Humanity?**

This chapter has focused prominently on Peter Singer’s duties of humanity, and sounded a decidedly pessimistic note about his approach. But an alternative vision of duties of humanity is available that might offer a more promising way of thinking about duties of assistance to reduce global poverty.

People living in poverty, by virtue of being human beings, having the capacity to suffer and sharing a fundamental interest in living a full, dignified life, deserve support from their fellow human beings. They are moral agents that are able to choose the life they want to lead among a set of acceptable options, and we all have duties to each other which cross national boundaries. One of those duties is to provide assistance when we can. Whether or not one has any association with one’s fellow humans who are living in poverty, we should help them to secure their basic human needs. This duty may not necessarily be discharged through donations to charity (though it may be) and it need not be insensitive to the social and political contexts in which poverty is produced. Based on a common humanity and nothing more, affluent people have duties to reduce the suffering associated with material deprivation, but these duties may involve derivative duties that ensure such efforts do not disrespect or harm their intended beneficiaries, and take seriously the agency of people living in poverty. This view may not convince people who deny such duties exist, but it is a more attractive formulation of the view that humanity has duties, while guarding against the worst abuses of the ‘savior’ mentality that can come to influence charitable giving.
Holding Psychopaths Responsible and the Guise of the Good
Abstract

Should we hold psychopaths morally responsible? Traditionally, this question is taken to turn upon whether psychopaths have the capacity for moral understanding. But what if psychopaths lack this capacity? There is a lively debate about whether psychopaths would nevertheless be responsible. This debate seems to take it for granted that this issue could be settled without further empirical investigation into psychopaths' deficits. Here I question this received orthodoxy. Central to my discussion is a distinction between two interpretations of "reasons for action." On a weak interpretation, one's reason for action is simply a consideration one takes as speaking in some way in favor of the action. On a strong interpretation, one's reason for action isn't merely a consideration that guides the action, but the consideration is also seen under the guise of the good. Whether psychopaths have the capacity to act for reasons in the strong sense seems to be a question that could be settled empirically. Indeed, there is some empirical evidence that suggests psychopaths lack this capacity. I argue that if psychopaths lack the capacity to act on reasons in this stronger sense, this will have a profound effect on the shape of the debate concerning their responsibility.

Much of the discussion of moral responsibility of psychopaths in the philosophical literature revolves around the question of the psychopaths' capacity for moral understanding, that is, understanding that there are nonstrategic reasons to comply with moral demands. Insofar as empirical evidence is relevant to this discussion, it appears straightforward that this would be evidence about various aspects of moral understanding found present or missing in those meeting the criteria for psychopathy, such as the grasp of the difference between moral and conventional prohibitions. I will take it for granted here that if it turned out that the psychopaths have the capacity for moral understanding, this would settle the question of their moral responsibility. In most plausible views, we would then have to hold them responsible. But what if, as it seems to me more likely, psychopaths lack the capacity for moral understanding? Much of the literature seems to assume that the question of psychopaths' responsibility would then be settled by a philosophical analysis, working out what the best theory of responsibility should say about the responsibility of those who lack the capacity for moral understanding. Of course, there is a lively debate, in fact a sharp disagreement, in the philosophical literature on whether lack of capacity for moral understanding excuses one from responsibility. Nonetheless, it appears that the issue could in principle be settled without further empirical investigation of the deficits of the psychopaths: what needs to be settled is the philosophical debate. I would like to question this received orthodoxy.
Here I can sketch only very roughly the two main approaches to thinking about the responsibility of those who lack the capacity for moral understanding. On the view prominently advocated by Gary Watson, there are two senses, or "faces," of deeming someone responsible.

In the weaker sense—Watson calls this the *attributability* sense—deeming someone responsible is, roughly, a matter of moral assessment; it involves appraising what the person is like as an agent. For example, we take someone to be morally responsible if he has exhibited a moral vice in his action. He has acted viciously, and we attribute a vice to him. I will return to Watson's view of what is involved in attributability. For now, suffice it to say that, for Watson, a psychopath who makes it his aim to cause someone pain and frustration clearly exhibits the vice of cruelty. In making this assessment of him, we attribute a vicious act to him and, in this sense, deem him responsible. And we can legitimately make this assessment of the action and thus attribute the action to the psychopath, irrespective of whether the psychopath has the capacity to understand that there are strong reasons not to act this way, or perhaps even irrespective of whether the psychopath has the capacity to understand that he has committed a morally prohibited or vicious act.

By contrast, holding someone responsible in the full sense—in what Watson calls the *accountability* sense—involves more than mere moral assessment or attribution of the action to the agent. When we hold people to account for their wrongdoing, we hold, among other things, that sanctioning or censuring them for a moral transgression would be warranted, at least prima facie. For Watson, this crucially involves making demands that the person recognize moral reasons, that is, the reasons not to do the wrongful act in question. Making such demands is senseless if the person lacks the capacity to recognize moral reasons. Since this demand is pointless, the lack of capacity for moral understanding directly implies the lack of responsibility understood as accountability.

On the contrasting view, associated most prominently with T. M. Scanlon and elaborated by Angela Smith, Matthew Talbert, and others, there is only one sense of holding someone morally responsible. So long as we can attribute an act to the person, we can also hold them accountable. One underlying rationale for this 'Scanlonian' view flows from thinking about why we hold people morally accountable. Specifically, the argument goes something like this:

- A key reason for holding someone morally and legally accountable for wrongful acts (that is, for censuring them) is to stand up against and protest a false or objectionable view of the moral standing of others that the perpetrator has expressed through her actions.
- A person can therefore be appropriately held morally accountable if she has in fact expressed, through her action, a false or objectionable view of how others
can be treated.

- According to the Scanlonians, a grasp of the concept of moral standing or of moral reasons is not necessary in order to express such an objectionable view. So long as a person is capable of acting on reasons, her actions do express a view of what the appropriate reasons for action are in the circumstances at hand. An action that altogether ignores moral reasons expresses the objectionable view that moral considerations are irrelevant; that the interests of the victim can be disregarded in one’s action. A perpetrator can express such an objectionable view even if she is incapable of understanding moral reasons.

- So, on this approach, we are entitled to hold psychopaths morally accountable so long as they are capable of acting on reasons. Under this condition their actions can express an objectionable view of reasons and we hold them morally accountable as a way of opposing or protesting against this objectionable view.

Note that, on this Scanlonian view, in holding a person morally accountable we do not necessarily demand that the person recognize moral reasons, so holding responsible a person who lacks the capacity for moral understanding is not pointless. The point is rather to stand up against and protest a false view of the moral standing of the victim.

If this Scanlonian view has merit, then to establish the moral accountability of psychopaths we need not worry whether they are capable of moral understanding. The question of their moral accountability rests on whether they are capable of acting on reasons. While this is an empirical issue, as far as I know, all parties to the philosophical debate accept it as uncontroversial that psychopaths are capable of acting on reasons. Some simply assert it. Watson, for example, states outright: “a [general] capacity to the effect that such and such is or is not a reason for action . . . psychopaths undoubtedly possess”. Others give examples of reasons for acting psychopaths presumably employ, such as, “I feel like a beer; this is a reason to go rob the 7–11”. Or take Talbert’s more sophisticated case of a malicious psychopath “seeing the prospect of another’s harm and suffering as a reason to act”.

However, I worry that this easy acceptance of the idea that psychopaths act on reasons conflates two rather different senses of the term “reason for action” There is an important ambiguity in what it means to act ‘for a reason’ overlooked in the debate about holding psychopaths responsible. The term “reason for action” admits of a weaker and stronger interpretation. In the weaker sense, one’s reason for action is simply a consideration one takes as speaking in some way in favor of the action. One acts on a reason in this sense when one pursues the action in light of this consideration; the consideration guides the action; one does what one does because of this consideration.
By contrast, in the more demanding sense, one’s reason for action is not merely a consideration that guides the action, but the consideration is also seen under the evaluative guise or the guise of the good. In acting on a reason in this sense, one evaluates the action as having a good-making feature. This importantly entails that in taking oneself to have a reason for action one believes that one is correct in doing so, that one is getting the matter right. Implicit here is the recognition of the possibility that one might not be correct after all—that in taking this consideration as a reason, one could be making a mistake. While a reason understood under the guise of the good is also a consideration seen as speaking in favor of the action, the reverse need not be the case: One can act on considerations that one takes to speak in favor of the action without seeing them under the guise of the good, that is, without the background notion that taking a consideration as a reason is something one can be correct or incorrect about.

This distinction is easily overlooked partly because, in paradigmatic cases of action, rational agents act on reasons in the stronger sense. The distinction comes into clear view in nonparadigmatic cases of human action, for example, actions of a very young child who is not yet capable of grasping (even implicitly) evaluative concepts and so cannot see anything under the guise of the good. The child may nonetheless see and act on all kinds of considerations in favor of the action: The child sees the anticipation that her parents will chase her and pay attention to her as reasons to run away; she sees the fact that a toy is new as a reason to play with it; she sees the fact that her sister has a pink hat as a reason to get one too; and so on. But, given the young age of children already capable of acting out of such considerations, there is a presumption against the notion that these children are evaluating any of these considerations as good ones to act on, that is, against the notion that they see themselves as being correct when they guide their actions by these considerations.

With this distinction in hand, let us return to the Scanlonian view of moral accountability. Recall that, for the Scanlonians, the decisive condition of moral accountability is that the person’s actions can express a view of what reasons for acting are *appropriate* in the circumstances at hand. These, I submit, must be reasons in the “guise of the good.” For if the person were merely capable of guiding her actions by certain considerations, without any implicit recognition that there is a correct and incorrect way of viewing the matter, then she wouldn’t be able to express any view of what reasons are appropriate; there would be no view of reasons here to oppose or protest. Of course, one could oppose the action (in the way that one may oppose the action of a toddler or even a lion), but one could not oppose her view of reasons because no view is expressed. In order to express a view, one must be in a position to declare its correctness, albeit perhaps only in some implicit fashion. But in taking considerations as reasons in merely the weaker sense, without the guise of the good,
one, by hypothesis, is not committed in any way to the correctness of the reasons—one simply happens to act on them—so one cannot implicitly declare the correctness or appropriateness of the reasons. The child in our example who sees the anticipated chase by an exasperated parent as a reason to run away is not thereby expressing the view that this is an appropriate way to treat the parent. Not being an evaluator yet, the child simply does not yet have any such views to express.

So, on the Scanlonian approach, the moral accountability of psychopaths must rest on whether they are capable of acting on reasons in the stronger, guise-of-the-good, sense. Note, however, that in accepting the seemingly uncontroversial assumption that psychopaths are capable of acting on reasons, all parties to the debate fail to attend to the guise of the good aspect. So they manage to establish only that psychopaths act on reasons in the weaker sense. The cited examples bear this out. The psychopath sees his desire for a beer as a reason to rob the 7–Eleven, but this doesn’t mean that he takes this reason to be appropriate or that he grasps that there is something here to be correct about. Similarly, the malicious psychopath sees the prospect of harming someone as a reason to act, but, for all we know, it couldn’t even occur to him to worry whether this way of seeing things is correct.

The recent empirical literature has not, to my knowledge, attended to whether psychopaths are capable of seeing reasons under the guise of the good, but there is some indication that they may not be. Cleckley appears to hold such a view in the passage cited by Nelkin (Chapter 3 in this volume): “The so-called psychopath holds no real viewpoint at all and is free of any sincere conviction in what might be called either good or evil.” The Psychopathy Checklist (Chapter 3, Appendix) includes items such as superficiality, “proneness to boredom,” and “lack of realistic long-term goals,” which are suggestive of a lack of evaluative commitments. Lack of regret is not on the Checklist but is often mentioned in the case studies of psychopaths; insofar as it is one of the marks of psychopathy, it would be a better indicator of a deficit in evaluative capacity. Nonetheless, it is not at all clear how to go about investigating this matter empirically.

If it turns out that psychopaths are capable of acting on reasons only in the weaker sense and have no understanding of the guise of the good at all, the Scanlonians would have to concede that it is inapt to hold the psychopaths morally accountable. According to the Scanlonians, the decisive condition of holding accountable is that the person’s actions express a view of what the appropriate reasons for acting are in the circumstances at hand. These are, as I have argued, reasons in the “guise of the good” sense. So if the psychopaths are incapable of understanding the guise of the good, they are not able to express the relevant view of reasons through their actions, and it would not make sense to hold them morally accountable.

But what implications would the same finding—that psychopaths are capable of acting on reasons only in the weaker sense and have no understanding of the guise
of the good—have for the rival Watsonian view of moral responsibility of psychopaths? Since Watson distinguishes two senses or faces of responsibility, we have to consider them in turn.

According to Watson, accountability, the stronger sense of responsibility associated with censure and sanctions, involves making a demand that the person recognize moral reasons. The application of this view to the psychopath would be unaffected by the finding that psychopaths have no grasp of the guise of the good, since, regardless of this finding, psychopaths are already thought to lack the capacities relevant for Watsonian accountability in virtue of being incapable of grasping moral reasons: The demand to recognize moral reasons cannot be sensibly addressed to them.

But what about Watson’s view of responsibility of psychopaths in the attributability sense? Here the finding that psychopaths are capable of acting on reasons only in the weaker sense and lack a grasp of the guise of the good might be relevant. In fact, the distinction between the two senses of acting on reasons brings out a seeming ambiguity in Watson’s understanding of attributability. In his earlier work on attributability, the “Two Faces of Responsibility” paper, Watson tied attributability to the expression of the agent’s true self and this in turn to the expression of the agent’s evaluative perspective. He wrote:

Aretaic evaluations thus differ significantly from other forms of appraisal. If I dance clumsily, it is inescapably true of me that I was (on that occasion) a clumsy dancer. But if what I do flows from my values and ends, there is a stronger sense in which my activities are inescapably my own: I am committed to them. As declarations of my adopted ends, they express what I’m about, my identity as an agent. They can be evaluated in distinctive ways (not just as welcome or unwelcome) because they themselves are exercises of my evaluative capacities.

(Watson 2004, 270–1; emphasis mine)

Thus, on this approach to attributability, if the psychopath doesn’t see reasons under the guise of the good, he doesn’t really have his own identity as an agent, so he has none to express. We cannot legitimately see him as responsible even in the weak sense of attributing the actions to him.

However, when Watson comes, in his later work, to explicitly apply his distinction between attributability and accountability to the case of psychopaths, he emphasizes different features of attributability. What’s at issue is whether the person exhibits virtue or vice in her action. Action that is in fact morally prohibited, such as hurting somebody, is, of course, not sufficient for this. The agent must have evinced disrespect for the victim. The story here is complex, but, at least in the paradigm case, Watson sees a psychopath who takes causing someone pain and frustration as her
reason to act to be exhibiting the vice of cruelty. What's now at issue for us is whether taking this consideration as a reason to act in the weaker sense is sufficient for vice. There is some indication in what Watson says that he would take the weaker sense of 'reason' to be sufficient. Watson emphasizes that the key feature of psychopaths' malice is that they "set themselves directly against others' aims as such . . . . Part of their intent in cruelty is to subordinate others' agency to theirs, in a way that seeks the others' recognition of their ability to do just this". It seems perfectly possible to have this complex intent and to take the corresponding considerations as reasons without seeing them under the guise of the good. What this complex intent is supposed to show is that the psychopath is 'into' her own agency in domination and thereby takes ownership of her agency. This distinguishes her from a mere animal and allows for attributability. But the psychopath can be 'into' her own agency in domination and thereby take ownership of her agency, without seeing her reasons under the guise of the good. As Watson acknowledges, children are capable of this kind of cruelty, and I would add that even children who lack the evaluative capacities required for acting under the guise of the good are capable of this kind of cruelty.

I am inclined to take the second approach to reflect Watson's more settled view of attributability. After all, in a paper that addresses the nature of psychopathic agency apart from the concern with responsibility, "Psychopathic Agency and Prudential Deficits," Watson argues that psychopaths' lack of ability to evaluate would furnish a good explanation for the comorbidity of moral and prudential deficit characteristic of psychopathy. So Watson seems to believe that psychopaths lack the ability to evaluate and thus also to see reasons under the guise of the good. Since he also thinks that psychopaths are responsible for their immoral actions in the attributability sense, the consistent view would be that attributability does not require seeing reasons under the guise of the good.

Overall, then, where does my investigation leave the debate about the moral responsibility of psychopaths? Initially, you might have thought that the question turned on whether psychopaths have the capacity for moral understanding. However, I invited you to take seriously the Scanlonian challenge to the view that lack of moral understanding is sufficient for lack of responsibility. The way this challenge is usually presented, it appears that so long as the challenge is philosophically sound, it would render psychopaths morally responsible. By contrast, I argued that, when this challenge is properly understood, it hinges the moral responsibility of psychopaths on whether they are capable of seeing reasons under the guise of the good. And this is an unsettled empirical question. Suppose we now set out to answer it. How would this shape the debate?

If it turns out that psychopaths are capable of acting on guise-of-the-good type of reasons, things are as the debate between the Watsonians and the Scanlonians
assumes—although it currently assumes it based on misinterpreted evidence—and the stalemate in the debate remains. If, however, it turns out that psychopaths are only capable of acting on reasons in the weaker sense and have no understanding of the guise of the good at all, the philosophical debate about their responsibility in the accountability sense can be resolved: The Scanlonians can be brought around to the Watsonian position. Recall that, according to the Scanlonians, the decisive condition of accountability is that the person's actions express a view of what the appropriate reasons for acting are in the circumstances at hand. These are, as I have argued, reasons in the guise-of-the-good sense. So if the psychopaths are not capable of understanding the guise of the good, they would not be able to express the relevant view of reasons through their actions, and they would not be appropriately held accountable. Thus both seeming contentious approaches to the accountability of psychopaths would give the same answer, and the question could be settled.

What we have also discovered is that the Watsonians and the Scanlonians differ in their accounts of attributability and that, perhaps surprisingly, they give conflicting answers about whether psychopaths are responsible in the attributability sense, if it turns out that the psychopaths are not capable of seeing reasons under the guise of the good. Since the Scanlonians collapse accountability and attributability together, once they back off on the accountability of psychopaths, they would then also have to say that the psychopaths cannot be deemed responsible in the attributability sense. By contrast, the Watsonians might have the resources to say that actions are attributable to psychopaths even if they don't see reasons under the guise of the good. Once you separate accountability from attributability, the conditions for attributability can be weak, and a mere explicit sense of one's own agency in dominance over another may be sufficient.

Even if you don't agree with these specific conclusions, I hope to have shown that the question of whether psychopaths are capable of understanding the guise of the good is very relevant to settling the responsibility of psychopaths and that it merits further investigation.
Diachronic Identity and the Moral Self
1. Introduction

Everyone undergoes many changes during the lifespan, and, in some cases, these changes may be quite dramatic. Intuitively, however, some of these changes do not threaten personal identity, while others do. We do not tend to suppose that identity is threatened by the physical aging process: wrinkles, gray hair, and the like. Significant psychological changes, brought on by aging, injury, or life events, seem to pose more of a threat. This, at least, is a thesis shared by many philosophers who have weighed in on debates about “diachronic personal identity” – the question, What kinds of continuity are required to qualify as the same person over time? For example, the most widely discussed suggestion is that diachronic identity is secured by the preservation of memory links of some kind. Others have emphasized narrative coherence and agency. Curiously, these competing proposals emphasize aspects of the mind that are not strongly associated with sociality. Admittedly, memories, life stories, and our capacity for agency all tend to involve social interactions as a matter of fact, but sociality is not constitutive of any of them. This reflects an individualistic orientation in leading approaches to diachronic identity, and we aim to offer a corrective. We believe that one of the most important aspects of diachronic identity is fundamentally related to social attitudes and social behavior: our moral values. We will defend the thesis that moral continuity (i.e., retaining the same moral values over time) is central to ordinary beliefs about what makes someone qualify as the same person as they advance through life. In fact, moral continuity is more important, according to our ordinary understanding, than memory, narrative, or agency. Each contributes to our sense of identity over time, but moral continuity contributes appreciably more. Our defense of this thesis is empirical. We present a series of studies that indicate moral continuity is regarded as extremely important for personal identity, and, indeed, more important than factors emphasized by prevailing theories. First, some preliminaries.

In saying that moral continuity is more important than other factors, we don’t mean to imply that it is the one true theory of personal identity and other theories are false. We don’t think the question of identity over time depends on some deep metaphysical fact. That is not to say we don’t think the question is metaphysical. We think it is metaphysical, but not deep; that is, it doesn’t depend on some hidden fact about the structure of reality. Rather, it depends on us. Facts about identity are a consequence of classificatory attitudes and practices. For any object that changes, those that label the object can decide which transformations matter for the application of that label. When it comes to persons, we suspect that multiple factors matter to us. We also suspect that different individuals and cultures may vary in which factors matter most. Here our goal is to show that randomly sampled members of our contemporary
Western society regard moral continuity as extremely important. This is part of our folk theory of identity. That theory does not aim to capture some deep underlying reality, but rather determines the reality. Folk intuitions establish the conditions of identity, and thus polling these intuitions is directly relevant to settling questions about diachronic identity. Unsurprisingly those intuitions point to multiple factors. We are content to show that morality is one of these factors and it outweighs some others that have received more attention.

If we are right that questions of personal identity are settled by how we do, in fact, classify, then this is a case where experimental philosophy can actually contribute to metaphysical debates. Surveys, in this case, do not just tell us what ordinary people think; they reveal the actual correlates of identity, because ordinary practices of classification determine conditions of identity. For those who are disinclined to share this view about the metaphysics of identity, the use of surveys may seem unmotivated. To those readers, we are content to point out that ordinary beliefs about identity are interesting even if they don’t cut metaphysical ice. Such beliefs may be important in human life, as when we deal with aging family members whose psychological profiles have changed, or with old friends who change political outlooks, or felons who undergo moral reform, or community members who become radicalized. Here we will approach the question of morality and identity at a high level of abstraction, using simply hypothetical cases, but the intuitions that we are tapping into may be important for many real-world decisions about how to regard and treat those who undergo evaluative transformations. We will broach these issues as metaphysical questions, as has been the custom in philosophical discussions of personal identity, but these practical issues may ultimately be more important.

Understood as a question about our classificatory attitudes and practices, it is probably best to resist the idea that personal identity is a matter of numerical identity in a mathematical sense. We doubt that ordinary folk would be much troubled by paradoxes that arise when philosophers imagine elaborate cases of fission and fusion. A broken teleporter, for example, might produce a duplicate of a Captain Kirk on the face of some planet without destroying the Captain Kirk who entered the device on the *Starship Enterprise*; the two individuals cannot both be numerically identical to the first, since they are not identical to each other, and numerical identity is transitive. What matters here is what Parfit calls survival. When we discuss diachronic identity, we mean to address folk intuitions about psychological transformations that threaten survival.

Our focus on diachronic identity should not be taken to imply that moral values matter only for survival, however. They presumably also matter for synchronic identity. “Synchronic identity” refers to those standing traits that are taken to make us who we are. Some traits, such as being allergic to peanuts, might not be crucial for a person’s identity, while others are. Often theories of diachronic identity and theories of
synchronic identity are specified in ways that are disjoint. Memory, for example, involves a link between present and past, and thus lends itself more to a theory of diachronic identity than synchronic identity. We think moral values are probably important for both. We are not the first to suggest that moral values matter synchronically. Frankfurt characterizes conditions of personhood in terms of there being a true self, and the true self is comprised of the set of volitional states that a person wants to possess. These presumably include moral attitudes, since moral attitudes are usually regarded as volitional, and most of us want to have at least some of the values that we possess. We will not take up the issue of synchronic identity here, but we do want to flag that it is plausible that moral values contribute to synchronic identity. If so, then the thesis that moral values contribute to diachronic identity promises to establish that at least some aspects of synchronic identity matter diachronically as well.

The thesis that morality is part of synchronic identity is not especially controversial, even if has not been given much explicit attention in the literature. The thesis that morality contributes to diachronic identity is likely to meet with more resistance. This may sound surprising at first, since it often suggested that diachronic identity is importantly related to morality. In his seminal discussion, John Locke says that personal identity is a “forensic” concept, meaning it relates to the assignment of moral and legal praise and blame. We care about identity for moral reasons. But Locke does not thereby mean that moral values are aspects of identity. Rather, he is suggesting that we need a theory of identity in order to assign moral responsibility. One might express this by saying that identity is for morality, which is decidedly different from the thesis that morality is partially constitutive of identity. Locke offers a theory of identity in terms of memory and “same consciousness” that makes no explicit reference to moral continuity. Thus, for him, identity matters morally, but morality doesn’t matter to identity. Here we will not discuss issues of responsibility, but we will try to show that Locke missed out on an important link between morality and identity, emphasizing one direction of that dependency and not the other.

To summarize, we are neither aiming to establish a deep metaphysical truth about diachronic identity, nor are we supposing that there is just one true theory. But we do aim to show that prevailing theories have neglected an important aspect of identity. Our studies suggest that moral continuity matters for survival over time according to lay intuitions, and it matters more than other characteristics that have dominated philosophical debates. If lay intuitions can arbitrate conditions on identity, it follows that moral continuity is an important part of what makes us qualify as the same people as we undergo changes in our lives. Thus, personal identity is more connected to our social nature than is often appreciated. We turn now to our studies.
2. Study 1: keeping promises

The idea that moral values might be central to how people conceive of their identity is neglected in most philosophical discussions, but it is passingly anticipated in a case devised by Parfit. Parfit helped reinvigorate debates about personal identity when he advanced the view that survival depends on psychological connectedness. When applied to non-human animals, survival may just be a matter of bodily continuity, but, for us, Parfit thinks that survival is a matter of having psychological continuity, which is based on "connectedness" – direct psychological links from one's present self to one's past. Following Locke, Parfit mostly emphasizes memory connections (or, for technical reasons, "quasi-memory" connections, because the term "memory" may presuppose identity). Officially, however, his account also includes other psychological connections. In this context, he mentions intentions, desires, beliefs, and character. These aspects of connectedness are given little attention, and moral connectedness is not explicitly discussed except in one example, Parfit's case of the Nineteenth Century Russian:

In several years, a young Russian will inherit vast estates. Because he has socialist ideals, he intends, now, to give the land to the peasants. But he knows that in time his ideals may fade. To guard against this possibility, he does two things. He first signs a legal document, which will automatically give away the land, and which can be revoked only with his wife's consent. He then says to his wife, 'Promise me that, if I ever change my mind, and ask you to revoke this document, you will not consent.' He adds, 'I regard my ideals as essential to me. If I lose these ideals, I want you to think that I cease to exist. I want you to regard your husband then, not as me, the man who asks you for this promise, but only as his corrupted later self. Promise me that you would not do what he asks.'

(Parfit 1984, 327)

As Parfit notes, the Russian's request, phrased in terms of successive selves, seems perfectly natural. We think that Parfit has hit upon an important link between values and identity, but he does not use the case to draw any general conclusions along those lines. Rather, it is used as a springboard to reflect on the idea that changes in identity can impact obligations, such as promises. Parfit does not dwell on the special role that moral values have in survival, much less on the possibility that they may be more important than memory – the aspect of connectedness that he emphasizes most in his other examples.

Our first study was inspired by Parfit's Russian. We sought to experimentally explore how intuitions about changes in values compare to intuitions about changes in memory. Participants were either given a case in which the protagonist has amnesia or a case in which the protagonist has a radical change in values. The amnesia case was
the following:

Consider this story about a Brazilian couple named Alberto and Claudia. Throughout their marriage, Alberto worked very hard to promote a political party called the Trabalhadores. He volunteered for the party and donated money. He also wrote a will asking Claudia to donate $1000 to the Trabalhadores after his death. Shortly after he retired, Alberto fell down and suffered a brain injury. The brain injury caused Alberto to lose his memories, but his values and his cognitive abilities, such as language and thinking, remained intact. After the brain injury, Alberto forgot all about the Trabalhadores, whom he had once supported. He also drafted a new will, which offered no money to the Trabalhadores. The next year, Alberto died. After Alberto's death, Claudia found both the old will and the new will, and her attorney told her that there is no legal way to determine which is binding, because Alberto did not explicitly nullify the old will. It is now up to Claudia to decide whether to donate $1000 to the Trabalhadores. She wants to follow Alberto's wishes, but she must decide which wishes to follow.

The values case went like this:

Consider this story about a Brazilian couple named Alberto and Claudia. Throughout their marriage, Alberto worked very hard to promote a political party called the Trabalhadores. He volunteered for the party and donated money. He also wrote a will asking Claudia to donate $1000 to the Trabalhadores after his death. Shortly after he retired, Alberto fell down and suffered a brain injury. The brain injury caused a complete transformation of Alberto's moral values, but his cognitive abilities, such as language and thinking, remained intact. After the brain injury, Alberto stopped caring about the Trabalhadores, whom he once supported. He also drafted a new will, which offered no money to the Trabalhadores. The next year, Alberto died. After Alberto's death, Claudia found both the old will and the new will, and her attorney told her that there is no legal way to determine which is binding, because Alberto did not explicitly nullify the old will. It is now up to Claudia to decide whether to donate $1000 to the Trabalhadores. She wants to follow Alberto's wishes, but she must decide which wishes to follow.

For each case, participants were asked whether Claudia should disregard or follow the newer will, which does not donate to the Trabalhadores. In the amnesia case, participants tended to say that Claudia should disregard the new will ($M = 3.09$ on a 6-point scale where 1 = completely disregard and 6 = completely follow). In the values case, participants tended to say that Claudia should follow the new will ($M = 4.9$). The difference between responses to these cases was significant ($t(41) = 3.34, p = 0.002$).
These results suggest that memory and morality impact intuitions about identity in different ways. If so, Parfit may not be entitled to treat them as equivalent. Their contributions to identity may differ, at least quantitatively (we will not explore qualitative differences here). This is an important first step in motivating our investigation into moral identity. At the same time, this study has a major limitation. The findings can be interpreted in two very different ways. On the one hand, one might think that our findings prove that memory continuity is more important for identity than moral continuity. On this interpretation, Claudia disregards her husband's revised will in the memory case, because she thinks he is no longer the same person after memory loss, and she feels a strong sense of obligation to his earlier self. In the case of moral change, she regards her husband as the same person, so she thinks he has the right to change the content of his will. If this is the correct interpretation, it would deliver a serious blow to the idea that moral changes impact identity. On the other hand, there is a reading that has the opposite implications. In the memory case, Claudia may reason that her forgetful husband is the same person as he always was, and she should honor his original preferences because after all they are still part of who he is; he has simply forgotten that. In the moral case, on this reading, Claudia regards her husband as a new man after the injury, and since the old Alberto is gone, Claudia has no more obligations to him, and should honor the will of her new husband.

Given these two opposing interpretations, we set out to test intuitions about moral change more directly. We decided to eliminate the issue of promise-keeping, which complicates Parfit’s original case and our variant. When a person makes a promise to someone who changes dramatically, two issues are easily confused: Is the change a change in identity? And, if so, does the person who makes the promise have an obligation to the original person, or is that discharged by that person’s demise? Since we are primarily interested in the first question, we sought to bypass the second. Still, this study established something important: memory and morality lead to different results. Our first goal was to look into this difference more deeply.

3. Studies 2a and 2b: moral self vs. mnemonic self

As noted, the previous study doesn’t establish whether people think that the changes to values disrupted the personal identity of the protagonist. It is open to interpretations on which moral change leaves identity intact. To address this directly, we designed a new vignette. We also wanted to verify that intuitions regarding moral change differ from intuitions regarding memory change, so we paired our new moral vignette with one involving memory. We asked participants about personal identity either in a case of amnesia or in a case of a change in values. The cases that we devised go as follows:
**Amnesia**

Imagine that John accidentally falls while walking in the mountains. The accident causes a head injury that has a profound effect on John’s memory. Because of the injury, John can no longer remember anything about his life before the accident. He is, however, able to form new memories. And his personality and values remain the same as before the accident. For example, before the accident he used to do helpful things for people in his community, and he continues to do this with enthusiasm after the accident, though he doesn’t remember that he used to do it before.

**Bad values**

Imagine that John accidentally falls while walking in the mountains. The accident causes a head injury that has a profound effect on John’s values. His memory and general intelligence remain the same as before the accident, but the injury causes John to stop behaving morally. For example, before the accident he used to do helpful things for people in his community, and, after the accident, he stops caring about any of that and only acts to fulfill his own happiness even at the expense of others.

In our first study with these vignettes (Study 2a), participants were asked to read one of these two scenarios, and then rate, on a 1–6 scale, the extent to which John is the same person as before the accident (1 = not at all the same person; 6 = completely the same person). We found that participants in the amnesia condition tended to think that John was the same person ($M = 4.32$; one sample $t(18) = 3.2, p = 0.0049$), whereas participants in the values condition tended to think that John wasn’t the same person ($M = 2.56$; one sample $t(17) = 2.9, p = 0.01$), and the difference between these responses was significant ($t(35) = 4.2867, p < 0.0001$).

These findings suggest two things: moral continuity is important to identity, on lay intuitions, and significantly more important than continuity of memory. There is, however, a small worry. In the values vignette, John becomes a morally worse person. Perhaps participants were somehow inclined to see John as transformed because they didn’t like his new values. Perhaps their judgments reflect negative attitudes towards the transformed John more than any reliable intuition about his identity. To determine whether this negative change is crucial to getting the effect, we also included a vignette in which John changes for the morally better:

**Good values**

Imagine that John accidentally falls while walking in the mountains. The accident causes a head injury that has a profound effect on John’s values. His memory and general intelligence remain the same as before the accident, but the injury causes John to transform morally. For example, before the accident he used to be very
selfish, working to fulfill his happiness even at the expense of others. After the accident, he suddenly stops acting selfishly and instead does helpful things for people in his community.

Again, participants were asked about the extent to which John was the same person as before the accident. Participants again tended to say that John was not the same person ($M = 2.59$; one sample $t(38) = 4.27, p = 0.0001$), and this was significantly different from the memory condition ($t(56) = 4.88, p < 0.0001$). By contrast, it made no difference whether the moral changes were good or bad ($t(55) = 0.089, p = 0.93$).

Study 2a strongly suggests that moral changes are regarded as threats to identity, regardless of whether those changes are good or bad. Still, this study leaves another question about the generality of the effect unanswered. Perhaps the effect is specific to how we think about the identity of other people. To explore this, we designed new vignettes (Study 2b) by simply replacing “John” with “you” in the above vignettes. We still got the key effect on morality vs. memory. People tended to say that if the accident caused significant changes to memory, they would still be the same person ($M = 3.91, one sample t(21) = 1.22, p = 0.235$), but that if the accident caused changes to their morals, they would not be the same person ($M = 2.09, one sample t(21) = 5.21, p < 0.0001$), and these responses were significantly different ($t(42) = 4.22, p < 0.0001$). And once again, even for positive moral changes, people still tended to say that such moral changes would mean that they were a different person ($M = 2.39, one sample t(22) = 5.9, p = 0.0001$). As with the vignettes involving another person, the vignettes focused on self showed no difference based on whether the changes were good or bad ($t(43) = 0.83, p = 0.41$).

Finally, we compared self vs. other on all the paired vignettes. The short answer is that we found no differences depending on whether the question was about self or other. Whether the question was about bad moral changes ($t(38) = 1.1, p = 0.28, n.s.$), good moral changes, ($t(60) = 0.59, p = 0.55, n.s.$) or memory changes ($t(39) = 0.9432, p = 0.35, n.s.$), no difference emerged for the self–other contrasts. No matter what we did, moral changes dominate over memory changes. In summary, moral changes matter to identity, whether they are good or bad, yours or another’s, and they matter more than changes in memory.

Defenders of the memory approach may resist this last verdict. They might even complain that we are not justified to suppose that memory and morality are disjoint capacities. Consider, for example, the finding that moral deficits are sometimes comorbid with deficits related to memory. Gerrans has explored such links in his discussion of individuals with ventromedial prefrontal brain injuries. Such individuals have moral impairments – they behave inappropriately and unreliably – as well as impairments with “mental time travel” – the ability to plan for the future. Mental time travel, in turn, is closely related to memory; the ability to look into the future shares
mechanisms with the ability to revisit the past. Still, we think it is noteworthy that ventromedial patients do not seem to have profound problems with autobiographical recall. Their memory impairments do not seem to involve the kind of memory capacities implicated in memory-based theories of personal identity. What these comorbidity studies show most clearly is that morality is not just important for decisions here-and-now. It is also crucial for future planning. This adds support to the idea that morality matters for survival over time. Rather than rescuing the memory theory, it helps to make sense of the idea morality has equal claim to being a capacity with temporal significance.

We don’t mean to suggest that memory did not matter at all. Loss of memory has some impact on intuitions about identity. But the impact seems to be rather small. Changes in memory fell above the midpoint on our scale, which measures intuitions about sameness of person. Averaging across 2a and 2b, the mean was 4.12, where 6 indicates completely the same person.4 In contrast, moral changes resulted in a mean of 2.51, where 1 indicates not at all the same person. The scalar nature of these results suggest that identity is graded, and probably no single factor matters. But it does seem that morality matters a great deal, and memory matters only modestly. This is striking because memory has been emphasized more than any other psychological trait in the literature on diachronic identity. It is a key element in Locke’s seminal thought experiments, and is also given disproportionate attention by Parfit. Memory is also the basis of identity in an influential theory developed by Paul Grice, and plays a central role in Sydney Shoemaker’s work on the topic. If memory makes only a modest contribution to identity, according to lay intuitions, then its prominent position in philosophy may need to be re-evaluated. In contrast to memory, moral continuity may be more important, at least in contexts like those that our vignettes are tapping into. We now want to suggest that moral continuity may also be more important than any of the main competitors to the memory theory of identity as well.

4. Studies 3a and 3b: moral self vs. agentic self

The previous results suggest that morality is regarded as much more significant to personal identity than memory. In all of the studies, we looked at cases in which an injury caused a brain lesion, which in turn caused the psychological changes to the individual. An obvious feature of these cases is that the brain lesions were in no sense chosen. These were changes that were externally inflicted on the individual. Perhaps this is the key to why we find people saying that there is a change in identity. Such an interpretation would be suggested by defenders of Agency theories of personal identity. Agency theories of identity have been defended by authors in the Kantian tradition, such as Christine Korsgaard. According to such theories, personhood constitutively involves agency, and agency involves the ability to make choices. Agency is said to secure a kind of identity through time, because, in making choices, we determine future
outcomes, and those outcomes therefore count as ours. Transformations that are imposed on us without consent, as in many thought experiments involving memory loss, are threats to identity not because memories are taken away, as Parfit might suggest, but because we are not the authors of such changes. For the Agency theorist, changes in one's psychological characteristics would not disrupt personal identity in cases where the changes are chosen. Korsgaard expresses the point this way:

Where I change myself, the sort of continuity needed for identity may be preserved, even if I become very different. Where I am changed by wholly external forces, it is not. This is because the sort of continuity needed for what matters to me in my own personal identity essentially involves my agency.

In our previous vignettes, there was no choice involved. Thus, while our results pose a problem for a Memory theory of identity, they might be accommodated by Agency theories. Korsgaard could say the moral transformation in our vignettes was perceived as a blow to agency not because it was a moral change, but because it was externally induced.

To explore this issue we conducted another study (3a) using a vignette in which there is moral change that results from a decision by the agent:

**Decision**

Imagine that John accidentally falls while walking in the mountains. He is not injured in the accident in any way; his memory and general intelligence remain unchanged. But the accident causes John to rethink his life. John decides that life is too short to waste time worrying about acting morally. For example, before the accident he used to do helpful things for people in his community, and, after the accident, he stops caring about any of that and only acts to fulfill his own happiness even at the expense of others.

The Agency theory predicts that we should think that John is still the same person since he decided to make these changes in himself. By contrast, the Morality theory predicts that since moral change is part of the scenario, this case would generate significantly lower judgments of identity. That's exactly what we found. The mean response was 2.14 on our 6-point scale, and this was significantly lower than a midline response (one sample t(21) = 4.71, p = 0.0001). In addition to the new vignette, we also presented a new set of participants with the Amnesia vignette from study 2a. People thought that John was the same person to a greater extent in the Amnesia case (M = 3.69) than in the Decision case (t(49) = 4.33, p < 0.001).

So we found that even when a person makes a decision to become less moral, people still think that the moral change has a major effect on their personal identity. But perhaps agency makes a difference over and above the effect of morality that we
keep finding. To assess this, we compared responses to the Decision case with responses to the case of Bad Values from study 2a, drawn from a new set of participants. In the Bad Values case, the moral change was the product of a brain lesion, and the mean response for this case was 2.58; this is not significantly different from what we found for the Decision case ($t(46) = 1.1465, p = 0.26, \text{n.s.}$).

Our Decision case focuses on a case in which participants evaluate the identity of another person in a hypothetical case. For our next study (3b), we wanted to see whether framing the case in terms of the self would change the results. Following Kant, Korsgaard notes that human beings can be studied from two perspectives: as objects of theoretical understanding, as in the human sciences, and as agents. Perhaps the vignette uses in Study 3a encouraged a theoretical orientation. To address this possibility, we reasoned that ideas about agency may become more salient if participants were asked to imagine making choices themselves. As before, we simply replaced “John” with “you” in the Decision vignette, and we also ran the Amnesia and Bad Values cases from study 2b. In the Decision case, the mean response was 2.16 on our 6-point scale, and this was significantly lower than an average response (one sample $t(24) = 4.98, p < 0.0001$), suggesting that even when imagining themselves making the decision, they still regard the future self as a different person. Once again, we compared Amnesia ($M = 4.05$) vs. Decision ($M = 2.16$) and found a significant difference ($t(42) = 4.8, p < 0.001$). But, as with the cases that asked about John, in the Self versions, there was no significant difference between moral change as a result of a lesion ($M = 1.85$) and moral change as the result of decision ($t(43) = 0.8754, p = 0.39, \text{n.s.}$). We also compared responses to the Self and Other versions of the Decision case and found no difference ($t(45) = 0.06, p = 0.95, \text{n.s.}$).

These results cast doubt on the suggestion that ordinary people regard agency as the lynchpin to identity. For the Agency theorist, changes that we choose should keep our identities intact. This is not what we found. Participants in our studies perceived moral change as a threat to identity even when such change is chosen. This remained true when they were asked to imagine choosing such changes themselves. The link between morality and identity is robust across this dimension. Moral changes continue to matter regardless of whether they are internally or externally imposed.

Korsgaard might not be very moved by our results. First, she might reply that her notion of identity is not as thin as we’ve implied. Elsewhere, she defends the idea of practical identity, which includes various self-conceptions, including one’s gender, ethnicity, social roles, and religion. This notion looks both broad enough and thick enough to include moral identity as a proper part. Consequently, she might be sympathetic to our approach, and insist that her account shares our prediction that moral changes threaten identity. We would be happy to have Korsgaard as an ally here, but we suspect that she cannot follow us all the way down the road we are
recommending. Despite her recognition that self-concepts (including one’s moral commitments) are part of practical identity, she is firmly committed to the view that agency is the ultimate arbitrator when it comes to change. If we choose to change our values, by an act of deliberation, we qualify as the same person. Current self-concepts may furnish us with reasons when making such choices, but we can also transcend those concepts, for Korsgaard, by viewing ourselves as members of the human community, and thereby abandoning commitment inherited from some narrower social group. When this happens, Korsgaard would say we have retained our personal identity. Our data suggest that this is not fully consistent with folk intuitions.

This brings us to another way in which Korsgaard might try to resist our conclusions. She might insist that ordinary intuitions are irrelevant to settling questions of personal identity. Her case for the link between agency and identity stems from an analysis of personhood, and that, she might suppose, is a conceptual claim that holds regardless of whether people happen to notice. For present purposes, we don’t want to get embroiled in debates about methodology. We are content to establish that people in our sample don’t regard choice as a difference-maker, even if, for some conceptual reason, they should. We are not out to refute Agency theories, or any others, because we are content to accept that many things may matter for identity. Our main observation here is that moral continuity continues to matter to people, regardless of agency, and that suggests it is a robustly important aspect of ordinary beliefs about diachronic identity.

5. Studies 4a and 4b: moral self vs. narrative self

The final prominent theory of personal identity that we’d like to consider is the Narrative theory. The core idea of the view is that a narrative – a life story – is necessary for establishing diachronic personal identity. The view has been developed perhaps most effectively by Marya Schechtman, who encapsulates the view as follows: “a person creates his identity by forming an autobiographical narrative – a story of his life”. Schechtman then goes on to suggest that the capacity to construct life narratives distinguishes human beings from some other sentient and sapient creatures. Memory and Agency theories entail that a creature can have a diachronic identity in virtue of having the capacity to recall past episodes or make choices. For Schechtman, that is not enough. Diachronic identity is secured by a kind of story telling. If this capacity is unavailable – if an individual cannot construct a narrative of different pieces of the past – then identity is disrupted.

What would a Narrative theorist say about the cases we’ve already investigated? Note that we find people regard moral change as having much more significant effects on personal identity than memory change. One might suppose that a loss of memory would be a greater insult to one’s life story than a shift in moral values. But the Narrative theory seems so flexible that it’s a bit hard to say. In each scenario,
one can imagine the individual developing a life story that included the accident or the decision. Or, one could imagine the individual not developing a story that included the accident. This flexibility of the Narrative theory is a kind of liability, since a theory that can twist to accommodate cases so easily might seem to lack a sufficiently firm backbone to be a substantive theory.

There are thus theoretical concerns associated with the Narrative theory, but rather than relying solely on these theoretical concerns, we designed a new study (4a) with a vignette that stipulated that the protagonist lost the capacity for narrative. The vignette went as follows:

**Narrative**

Imagine that John accidentally falls while walking in the mountains. The accident causes a head injury that has a profound effect on John's ability to construct coherent stories. Before the injury, John often thought about his life in terms of an ongoing story. After the injury, he can no longer tell stories or follow stories, and he stopped thinking about his life in terms of an ongoing story. But his personality and values remain the same as before the accident. For example, before the accident he used to do helpful things for people in his community, and he continues to do this with enthusiasm after the accident, though he doesn't think about how this fits into his life as part of an ongoing story.

The vignette claims that John can no longer form a life narrative. The narrative theory of personal identity would thus hold that John's identity has been radically disrupted. But this is not how the participants saw it. Even though John's ability to form narratives was destroyed, participants still tended to say that he was the same person after the injury ($M = 4.11$; one sample $t(26) = 2.67, p = 0.0128$). For contrast, we also had another set of participants evaluate the Bad Values vignette (study 2a). As before, for that case, participants tended to say that John was a different person ($M = 1.89$; one sample $t(27) = 8.9, p < 0.0001$), and this differed significantly from responses to Narrative ($t(53) = 7.64, p < 0.0001$).

As in previous studies, we also wanted to see whether there was something special about using a third-person case, so we rendered Narrative into a first-person case by replacing “John” with “you”. The results were much the same. People still did not give the response predicted by the Narrative theory. Rather, responses were on the “same person” side of the scale ($M = 3.77$; though not, in this case, significantly different from midline: one sample $t(25) = 0.86, p = 0.39$, n.s.). Again, we also ran a first-person version of Bad Values and, as expected, participants tended to say that they would not be the same person in this scenario ($M = 2$; one sample $t(21) = 10.2, p < 0.0001$). Responses to the Bad Values case differed significantly from responses to the Narrative case ($t(46) = 4.8610, p < 0.0001$). In addition, when we compared the first-person
version of the Narrative case to the third-person version, we found no difference ($t(51) = 0.89, p = 0.3767$).

For these studies, we asked participants to explain their answers. Since the idea of narrative construction is a bit less familiar than memory or agency, we wanted to get a flavor of how people conceived it. Some subjects did cite the significance of narrative, e.g.:

"Telling stories is a main part about who you are. If you are unable to do that then I think that significantly alters parts of your communicative skills and may make you less outgoing."

"I think that narrative is the way we construct and understand our lives. Without this, a person would be completely different."

But more subjects registered an explicit rejection of the significance of narrativity for personal identity, e.g.:

"I think a person’s ability to tell a coherent story is a miniscule part of who they are, so I think John is basically the same person. John or the person he is telling a story to may experience some confusion but that doesn’t mean what essentially makes John himself has changed."

"Although he can no longer think about his life as an ongoing story or follow stories these are very minor parts of who he is. The major parts, his personality and values are still there and make up a very large part of who he is."

"He is not able to tell his story but still has the values of his life intact. He is still able to help his community and be a good person. He just cannot put things together as far as the story of his life."

For the moral transformation, by contrast, subjects tended to be quite explicit in endorsing the significance of morality for identity:

"Our morals are part of who we are as a person, and determines our interactions with others and social success. If he isn’t as moral, he isn’t the person that everyone remembers."

"If John’s values change he is not the same person. However, I believe the change is not his fault whether it be for the better or worse."

"While John may remember who he was, and may have the same level of intelligence, his moral code and personality is what made him who he was. If that changes then the person changes with it."

One subject actually articulated the Morality theory of personal identity with perfect economy:

"His morals changed. Therefore he changed."
Overall, the results provide no support for the idea that Narrative is central to how most people think of personal identity. More importantly, though, even if narrative matters somewhat to identity, our results suggest that the significance of narrative pales compared to the significance of morality.

6. Summary and implications

We have now presented a series of studies exploring folk intuitions about the relationship between morality and the self. Our first study took inspiration from Parfit and established that changes in values and changes in memory have a different impact on intuitions about spousal obligations. This initial finding motivated further investigation, and we then probed intuitions about survival more directly, asking participants whether someone is the same person after undergoing a moral change. We found strong support for a negative answer: moral change significantly alters personal identity. This was true whether the moral change was positive or negative and whether it was presented from a third- or first-person perspective. This suggests that moral continuity matters a great deal on folk understanding of diachronic identity.

We take this to be our main positive finding. In addition, we compared the importance of moral continuity to other psychological factors that have been emphasized more frequently in the philosophical literature. We identified three factors that figure prominently in some of the most prominent theories of personal identity: Memory, Agency, and Narrative. In each case, moral continuity was found to be more important, according to lay intuitions. Loss of memory and narrative capacity has only a modest impact on judged sameness of person, and the presence of agency did little to mitigate the insult to identity associated with moral change.

As we’ve stressed, we don’t mean to imply that these other approaches to personal identity are mistaken. We think there is no deep metaphysical fact about identity, and multiple factors may be used to individuate persons and track their identity over time. At the same time, our findings do indicate that morality may matter more than some factors that have dominated discussion. This suggests that the moral dimensions of identity deserve more study.

In making this comparative point, we are not trying to suggest that morality is the most important factor in identity. Other factors may be equally important. We have not looked at non-moral values, occupation, and personal relationships, for example. That said, moral continuity has been compared to other psychological traits and capacities in follow-up studies that were recently published. Participants were asked to imagine scenarios in which there were changes in personality and a variety of cognitive and perceptual faculties. All of these make some contribution to intuitions about survival, but none has the same impact as moral change. Thus, ongoing research done in support of our core findings here has yet to unearth any feature of the mind that is given as much weight in diachronic identity as morality.
It might be objected that folk intuitions cannot settle questions about what matters to identity. As noted in section 4, we allow that conceptual analyses of personhood could reveal conditions of survival that the folk haven’t fully appreciated. On the other hand, we think questions about personal identity can be directly illuminated by research of this kind. Often, when we ask about survival we are asking about what people really care about. Surveys can help us answer this question. Suppose a loved-one changes her values. Would we think of her as the same person? That is a question that can be investigated using survey methodology. It would be valuable to explore real-world cases as well, but surveys are well-suited for tapping into what we care about.

It’s a reasonable further question whether there are some other things that we should care about more. But if we do care about morality, then that fact alone suggests that morality plays a role in the individuation of persons. It is part of how we construct the category. It’s independently plausible that personhood is a category whose boundaries are determined by us, not a natural kind with a biologically given essence. If so, the criteria we use in person individuation are partially determinative of what identity consists in. In our sample, morality is clearly part of the story. Future work will have to explore whether other populations share these intuitions.

We want to end by considering some of the implications that might follow if moral continuity is indeed a crucial component of diachronic identity. First, though our vignettes are fanciful, moral changes do occur in many people’s lives. These include religious conversion experiences, indoctrination, radicalization, and shifts in political orientation. Each of these may suggest a kind of discontinuity or rupture in identity. This is not an affront to intuition. We do often feel alienated from those who undergo such changes. The work presented here sheds light on those feelings. We implicitly register that those whose values change can cease to be fully recognizable as the same persons, despite many other continuities. When family members try to recover a relative who has been radicalized, for example, there is a palpable sense that the old self must be brought back to life.

Moral identity may also have implications for cases of prison reform. It’s a familiar refrain from prisoners at parole hearings that they have changed – “I’m not the same person.” Here’s one recent example. Jonathan Coryell was convicted (with two others) in the 2001 murder of Jeff Smulick. Coryell was sentenced to 22 years. In 2010, Coryell tried to get a reduced sentence by emphasizing changes since sentencing – engaging in thousands of hours of community service and working with at-risk youth. Coryell stressed that he now wanted to devote his life to such children. At a key point in his request, he writes, “I am not the same person who this court sentenced years ago.” Our findings suggest that such claims are corroborated by folk intuition. Morality is regarded as an important dimension of identity, and thus cases of reform – or
corruption – may qualify as changes of identity. If so, this may have implications for exculpation. We are currently investigating the impact of reform on attitudes towards parole. From a descriptive and normative perspective, moral change may imply that certain inmates should be given reduced sentences, not just because they are less likely to offend after reform, but because they are not the ones (to some degree) who committed the earlier crimes.

There is another implication that also bears on responsibility in criminal contexts. Certain individuals may have an impaired capacity for morality. If so, that dimension of personhood, and personal continuity over time, will be correlative diminished. Consider clinical psychopaths. According to some researchers, psychopaths cannot fully grasp moral rules. They regard moral violations as akin to etiquette violations and cannot fully articulate why it is wrong to cause harm. This raises questions about their culpability. Legal ethicists ask whether those who cannot fully grasp morality can be held morally responsible. The concept of moral identity introduces further questions that have received less attention. Psychopaths may be said to lack moral identity, both synchronically and diachronically. If so, we can ask whether psychopaths qualify as persons in the same sense as the rest of the population and whether the conditions for continuity in their lives differ from ours. In ongoing work, we are looking at how individuals with psychopathic tendencies view the role of morality in identity. It will be equally important to explore how we view them in this connection. One possibility, for example, is that certain moral relations that have a temporal dimension will work out differently for this population. Would a promise made to a psychopath be binding to the same degree and for the same reason as a promise made to a non-psychopath, if they lack moral identity?

Psychopaths are not the only individuals who have been alleged to lack a full moral capacity. The same claim is sometimes made about young children, though they ordinarily acquire moral values later on. It is an implication of our view that young children lack an aspect of identity that is regarded as important in adults. That means that the emergence of morality in early life brings a kind of identity into place that may not be there, at least to the same degree, at the outset. It would follow that there is a sense in which a child is not the same person before and after moral education. This may sound odd at first, but it is not an objection to the theory. First of all, other aspects of identity may be in place, so we need not conclude that there is no diachronic continuity here. Second of all, it is not outlandish to suppose that aspects of the self emerge in early life, and that we are not fully identical to our premoral counterparts.

A slightly more embarrassing implication concerns moral improvement. Suppose one takes it upon oneself to embark on a program of moral enhancement. Realistic cases include those who discover that they harbor forms of bigotry and set out change that. Such self-directed consciousness-raising seems like a good thing, but, if
morality is part of identity, it may qualify as a form of suicide. In improving our values, we may be destroying our old selves. Given other aspects of continuity, this may not involve a complete loss of self, but the ideas we have been exploring suggest that moral improvement is a kind of self-inflicted harm. That may sound problematic, not least of all because it undercuts the incentive to improve ourselves. We think the worry can be addressed. The fact that one has to sacrifice aspects of identity in moral reform does not render such efforts odious. Instead, we can think of moral improvement as a justified self-sacrifice: I will give up my old self for the sake of a better self. This does not strike us as an odd way of talking. Those who shed deeply entrenched forms of bigotry may come to regard their past selves as alien, and, before embarking on an effort to change, the old self may justly assert: I don’t like myself; I want to become a different person. Of course, we may feel ashamed of our old selves, but shame can also extend to close kin. Looking back, we can rejoice in the fact that our flawed predecessors made the ultimate sacrifice on our behalf.

A further implication concerns the question of how many personal identities a human being is likely to have in a lifespan. In advancing his neo-Lockean view, Parfit provokes readers by suggesting that, if we focus on connectedness when evaluation personal identity, persons are relatively ephemeral; human lives lack reliable memory connections over time, and this implies that the self can be rather ephemeral. One might think that an even more radical instability follows from the idea that identity involves morality. Moral values can be likened to character, and there is a sizeable literature suggesting that people lack robust, causally efficacious, temporally stable character traits. Does the view that morality is part of identity therefore entail that identity is fleeting or fragile? We don’t think so. Unlike some aspects of moral character, which may lack stability (e.g., honesty), we suspect that moral identity is relatively stable across the lifespan. There is evidence, for example, that political party affiliation remains remarkable stable. Sears and Funk found that the correlation in party membership for an American sample was 0.80 over ten-year periods, and close to 0.70 when measured over almost four decades. Skeptics about character sometimes cite a study showing that students at Bennington became more liberal during their years at the college, indicating a high degree of moral plasticity. This we don’t deny, but follow-up work showed that such changes are long lasting; fifty years on, the Bennington graduates remained more liberal and more political than peers who studied elsewhere. Given that political values have a strong moral dimension, we can infer that moral values can be stable for long periods. That doesn’t mean there is no change: we have given examples of conversion experiences, and studies suggest that people become slightly more conservative as they age. But moral values may be more stable than, say, memory and narratives.

The reasons for this stability deserve to be explored. One explanation is that
moral values are grounded in emotions, and thus threats to moral values are upsetting to us. Mechanisms such as cognitive dissonance may make us resilient to moral change. Another compatible explanation is even more interesting. Moral values may remain stable because they are socially sustained. When one joins a moral group, such as a political party, one ends up with peers who serve to regulate attitudes and behavior. Group members are our conversation partners, and they penalize us when our beliefs stray. To change values would threaten close social relations; we risk scorn, ridicule, and even ostracism. Through both positive and negative reinforcement, social groups enforce attitudinal continuity. This relates to an idea defended by Maria Merritt; though skeptical about character traits, she argues that social relationships can serve to impose a kind of behavioral consistency. In the case of memory, such mechanisms are less likely to be in place.

This brings us to a final implication – one which bears most directly on the central theme of this volume. As noted at the outset, many of the prevailing theories of diachronic identity are individualistic in orientation: Memory, Autonomy, and Narrative are all things that an individual could, in principle, have if she were alone in the world. Memory and narrative aspects of identity are also individualistic in a further way: they are things that differentiate us. Each person has her or his own biography – a unique personal narrative. Autonomy is a feature that all healthy human beings share, but it is a generic aspect of human nature, not something that we share in virtue of being linked together in social groups (except on certain Hegelian perspectives, which we won’t review here). Morality is different. Unlike memory and narrative, moralities are shared. They are systems of norms that we have in common with other people in our communities. They are also socially inculcated, and moral rules concern behavior and are intended to coordinate social interactions. Moral beliefs can be idiosyncratic, but the very function of morality is social. In these respects the moral aspects of our identity are collectivist, rather than individualistic. To state it ironically, it is a part of personal identity that is not personal, in the narrow sense of being distinctive of individual persons.

The claim that personal identity is social can be understood as a thesis about folk beliefs; we regard morality as important to identity; morality tends to depend on membership in social groups; therefore, we are implicitly committed to a social view of the self. At the beginning of this discussion, we suggested that folk beliefs determine metaphysical facts when it comes to personal identity. If so, the claim that personal identity is social can be read as a metaphysical hypothesis. What we are as individual persons is both influenced by social factors and sustained by social affiliations. We come into our personhood, one might say, by forging links to moral communities. What makes me me is, in part, something that binds me to you. Though this has not been our focus here, we think it is perhaps the most important implication of the view we are
advocating. In overlooking the importance of morality to personal identity, philosophers have also been neglecting a way in which human selfhood is social, and the social affiliations inherent in our moral values secure a kind of stability that goes beyond security and camaraderie. In some sense, having a life – persisting through time – depends on traits that can be described as social.
CHAPTER 5

Human Genetic Engineering
Of growing concern to bioethics are some recent innovations in genetics and biotechnology that seem, from a moral perspective, a two-edged sword. Their implementation in diagnosing genetic disorders strikes us as beneficial, while raising worries about the possibility of causing harm to present or future people. As we'll see in this chapter, their use in human reproduction creates moral problems of its own, especially in connection with prospective human cloning and the altering of genes for human enhancement.

15.1 Designing Deaf Baby

Sharon Duchesneau and Candy McCullough, a deaf couple in their 30s from Bethesda, Maryland, had already begun building a family when in 2002 they sought reproductive assistance in order to conceive a deaf child. With their first child, they were denied a chance to select for implantation only embryos with the trait of deafness that they wanted. So they sought in-vitro fertilization with sperm from a friend who had five generations of deafness in his family. They eventually succeeded in having Jehanne, a deaf girl they intended to bring up in the deaf community and culture. In their view, deafness is not a disability but a linguistic trait and cultural identity to be embraced with pride, something akin to being Black, East Asian, or Hispanic.

Like many in the deaf community, Duchesneau and McCullough cherished the use of sign language and the complex culture around it, which they wanted to share with their children. Furthermore, they thought they would be better parents to a deaf child. At seven months through the pregnancy, Ms. Duchesneau told reporters: “It would be nice to have a deaf child who is the same as us. I think that would be a wonderful experience. You know, if we can have that chance, why not take it? A hearing baby would be a blessing. A deaf baby would be a special blessing.” Since deafness is one of the many traits that recent reproductive technologies allow to select before implantation, she regarded as a prejudice the health care providers’ refusal to help them select embryos with deafness. As it happens, their son, Gauvin, was born with partial hearing in his right ear. The couple declined medical advice to fit him with a hearing aid so he could develop comprehension of speech. They have also declined that advice for Jehanne. In their view, it should be up to the children to make that choice when they get older.

***

Duchesneau and McCullough decided to enlist a deaf sperm donor because they were denied PGD, the procedure outlined in Box 15.1. Although PGD was introduced to detect and discard embryos with certain genetic disorders linked to sex, it soon became available for selecting embryos for nonsex-related medical conditions. In 2003, French
doctors learned how to diagnose dwarfism with PGD. They soon discovered that some couples were requesting PGD for the transferring of embryos with the dwarfism trait. PGD might also be used to select for the presence of traits such as sex, blindness, and some mental impairments. Yet, as in the deaf couple's case, most clinics refused to provide it for designing a disabled baby. Of 190 clinics surveyed in 2006 by Johns Hopkins University, 3% reported the use of PGD for selecting embryos with a disability trait.

**BOX 15.1 PREIMPLANTATION GENETIC DIAGNOSIS (PGD)**

- PGD, combined with in-vitro fertilization (IVF), is used on one or two cells of a 3-day-old embryo, which by then has between five and ten cells.
- The cells are screened for single-gene defects or for genetic mutations linked to heritable disorders.
- Embryos with disease traits are discarded, the others transferred for implantation to the woman's womb, stored for later use in reproduction, or donated for research.
- The first birth from a PGD-selected embryo occurred in 1990; by 2002 there were about 1,000 according to the PGD International Society.

Why might many think such use morally wrong? Dena Davis provides an Autonomy Argument:

1. Designing a disabled child imposes on her the parents' conception of a good life.
2. Limiting her choices violates her right to an open future.
3. Violating a child's right to an open future is morally wrong.

4. Therefore, designing a disabled child is morally wrong.

Davis thinks that such decisions narrow the scope of the future child's choices when she grows up, thus violating her right to an 'open future'. Among her disadvantages later in life will be fewer options in career, marriage, and other life goals. The Duchesneau and McCullough's decision illustrates premise (1) because it imposed the couple's conception of deaf culture on their offspring. For Davis, their rationale was analogous to that of parents who decide to deprive their children of an education or life-saving treatment on the basis of their religious or life-style preferences. If those decisions undermine the children's wellbeing, they would be wrong on nonmaleficence grounds. But for Davis they primarily violate respect for autonomy by treating children as mere means to their parents' goals. Her objection is independent of whether deafness is a disability or a cultural identity. Yet it entails that "deliberately creating a deaf child is a moral harm".
However, some utilitarians reject the Autonomy Argument because of its incompatibility with *procreative freedom*, which they value for the reasons outlined in Box 15.2. As you may recall from Chapter 14, libertarians also value it but not as means to something else. In either view, parents are at liberty to design a disabled child unless her disability is such that her life is not worth living. Since deafness is compatible with having a life worth living, it is morally permissible to design a deaf child. Here utilitarians and libertarians invoke their own intuitions about Derek Parfit’s Nonidentity Problem.

Greatly simplified, this is the problem of whether parents have wronged their offspring by conceiving them with disabilities compatible with a life worth living. For these utilitarians and libertarians, a child with a worth-living life cannot be harmed by being brought into existence because had the parents made a different procreative choice, that child would not exist. Fueling their intuitions are

1. Genetic essentialism—Different genetic origins necessarily result in different persons. In the deaf couple case, a different procreative decision would have produced either no child or a different child.
2. Person-affecting harm—In designing a disabled child, what matters is harm to the child or to others.
3. Reproductive freedom—Making self-regarding reproductive decisions is valuable either by itself (libertarian view) or as a means to the best life (utilitarian view).

For example, Julian Savulescu acknowledges that the deaf couple’s decision resulted in the birth of a disabled child. But since her life is worth living, it did not harm that child: “Couples should be free to request and obtain genetic testing provided there are sufficient resources and their choices do not harm the child produced or other people, even if this deliberately brings a child into the world with what most people judge to be worse than average prospects.” As Jonathan Glover puts it, how could the couple owe to that child that she was never born? Thus, a crucial part of the utilitarian/libertarian argument is:

1. Designing a disabled child is morally wrong only if the child’s life is not worth living.
2. The deaf child’s life *is* worth living.

3. Therefore, her parents’ choice was morally permissible.
Premise (2) is uncontroversial. But (1) is vulnerable, even after accepting the value of reproductive freedom and the failure of state-mandated policies restricting such freedom (see “The Early Eugenics” below). Dan Brock, for example, is committed to offer this *Wrongful-Disability Objection*: even if the deaf couple did not harm the child, they failed to bring about the best state of affairs—something morally wrong for utilitarianism. They thus produced a nonperson-affecting harm that outweighs the value of procreative freedom.

Of these three arguments about the morality of such parental choices, only one finds them permissible. But all three need further assessment since they rest on controversial moral principles or calculations and conceptions of the benefits and harms brought about by those choices. Although such assessment goes beyond our purpose here, an informed view on the matter requires a broader look at the uses of genetics in human reproduction.

15.2 Some Moral Controversies in Reprogenetics

Crucial to current developments in reprogentic, a combination of new reproductive technologies with genetic methods, was the introduction of *in-vitro fertilization* (IVF) in the early 1980s. Such combination made possible the development of preimplantation genetic diagnosis (PGD), originally designed to detect the presence of genetic medical conditions in IVF embryos before their transfer to the woman’s uterus, especially when parents are at high risk for genetic disease. But as we saw, the use of PGD has expanded considerably beyond its original purpose. Furthermore, genetics may make possible human reproductive cloning (HRC), to which many object on the moral grounds discussed in this section. Triggering their fear is the belief that these reproductive interventions amount to a new eugenics morally analogous to the much condemned, old eugenics of the early 20th century.

*Preimplantation Genetic Diagnosis*

Determining an embryo’s sex was initially part of the screening for specific sex-linked disorders such as hemophilia or Duchenne muscular dystrophy, but some parents now request PGD for selecting their offspring’s sex. Selecting embryos for the absence of disease traits falls under preventive medicine and is therefore less morally problematic than selecting for the presence of nonmedical traits, such as sex or eye color. As in the practice of using PGD for designing a disabled baby, there are reasons for and against using PGD for selecting embryos with certain nonmedical traits. For one thing, there is the above considered utilitarian/ libertarian argument, based on procreative freedom. Furthermore, given the following utilitarian principle, PGD to select for medical and nonmedical traits is not only morally permissible but obligatory:
Procreative Beneficence

“[Reproducers] should select the child, of the possible children they could have, who is expected to have the best life, or at least as good a life as the others, based on the relevant, available information.” (Savulescu, 2001: 415)

Since PGD maximizes the possibility of conceiving children with the best chances of having the best life, Procreative Beneficence makes PGD selection prima facie obligatory. But Julian Savulescu faces the problems of

1. Defining ‘best life’
2. Recommending the use of PGD to select embryos for the presence of traits such as race and sex.

In response to (1), he defines ‘best life’ as that of a child with traits that can bring her as many social advantages as possible. So his view is a type of perfectionism that runs into problem (2): it is well known that in our society, being of a certain race or sex is an advantage. Procreative Beneficence condones such biases. Here Savulescu responds that it is society that is responsible for eliminating those biases.

Some utilitarians similarly justify PGD for creating a savior sibling—i.e., a child conceived to save an existing sibling affected by a disease of the blood or immune system. The savior sibling may, for example, be the source of hematopoietic stem cells from the umbilical cord or bone marrow to save, or ameliorate, the condition of the other child. The birth of a savior sibling is particularly crucial in situations where no compatible cells are available for a transplant that is needed to save a child affected with Fanconi anemia, who faces the prospect of premature death unless she receives blood from the umbilical cord of a perfect match. With PGD, embryos that are a good genetic match can be selected for implantation. For critics of the practice, procreative freedom and beneficence are not enough to justify the wrongness of treating the savior as a mere means to save its sibling. For defenders, evidence would be needed that the practice can harm the savior, or that it precludes wanting and loving her for her own sake.

On the other hand, the wrongness of sex selection for nonmedical reasons (family ‘balancing’ or personal and cultural preferences) is often cited against PGD. But as shown in Box 15.3, PGD is not the only procedure available today for increasing the chances of having a child of a specific sex.

Of these three procedures, the least morally controversial is sperm sorting (because the selection is done before conception), and the most controversial is sex-selective abortion (because it is done after the embryo has implanted). All three are costly and face some of the moral objections to sex selection discussed in Chapter 12. Since PGD for sex selection faces similar objections and they were shown inconclusive, the moral wrongness of the practice remains unsettled (though it is illegal in Canada and the
United Kingdom). In addition, the moral wrongness of sex selection would not undermine PGD in general but only when applied to sex selection.

Another objection to PGD, the Expressivist Argument, contends that in allowing parents to discard IVF embryos with disability traits, the use of PGD implicitly devalues the lives of disabled people. It thus discriminates or expresses disrespect toward the ways of life of the disabled and treats them as something to be rejected or held in contempt. Those ways of life deserve respect, just as other ways of human life do.

The Expressivist Argument, if sound, justifies banning on moral grounds not only PGD but also abortion for fetal defects and end-of-life measures for impaired neonates. The best reason for its premise is that such procedures express a bias analogous to sexism, racism, and homophobia. This analogy does not object to all new reproductive technologies (or to all abortions). But it does imply that parents and providers who use PGD or prenatal screening for genetic disability are acting wrongly. That seems an overreaction, given how severe genetic diseases and disorders can be. For one thing, given society’s limitations, not every parent can care for a child with special needs. It seems that those who cannot do it are under no obligation to do it. After all, many ethicists agree that morality cannot be so demanding that saintly or heroic action is an obligation.

Furthermore, as we saw in Chapter 12, wrongful-life litigation suggests that some severe congenital impairments may render a life not worth living, as judged by the person whose life it is. In light of these replies, the Expressivist Argument seems compelling only in the sometimes-denounced scenarios where the parents have refused PGD because they can care for a baby with special needs, but the clinic,
concerned with its IVF success rates, insists on implementing it. Other critiques of PGD range from the Embryo-Mishandling Objection discussed in Chapter 14 to the Hubristic-Motivation Appeal considered below.

Human Reproductive Cloning

Some Facts

More than 100 years ago, scientists first demonstrated that simple organisms could be replicated by embryo-twinning at early stages of development. In 1958, John Gurdon performed the first successful transfer of a tadpole's intestinal cell nucleus into an enucleated frog egg, creating a clone or genetic replica of that tadpole. Known as 'somatic cell nuclear transfer,' this cloning technique proceeds by injecting the DNA of a donor cell, or fusing the entire cell, into an enucleated egg. Any resulting embryo is immediately implanted into a gestational carrier, in the hope that it will result in a live-born clone. In 1996, UK scientists Ian Wilmut and Keith Campbell transferred the nucleus of an adult sheep's udder cell to an empty egg to produce the first mammal cloned from an adult somatic cell. After 277 attempts that led to the creation of 27 sheep embryos and only 13 implantations, Dolly the Sheep was the only clone to survive. Success rates in animal cloning are disappointing, with primates posing the greatest technical challenges. But in 2013 scientists created human embryonic stem cells from a baby for the purpose of studying its rare genetic disorder. The difficulties facing such therapeutic cloning became evident during the 2004–05 scandal involving a South Korean scientist suspected of faking his data.

To date, there is no evidence that human reproductive cloning (HRC) has actually been done, despite isolated claims to the contrary. Although it is unlikely to happen soon, the fact that it may become scientifically possible has already raised moral and legal controversy. If animal reproductive cloning were to be perfected so that HRC could be tried without considerable risk of harm, should it be done?

The Morality of HRC

For bioliberals, who are enthusiastic about the use of genetics and biotechnology on humans, HRC is just another new reproductive technology. Its permissibility when it reaches adequate levels of safety follows from its benefits and the value of procreative freedom. Unlike critics who reduce HRC’s possible ‘benefits’ to creating a mini-me or replicating a celebrity, bioliberals hypothesize that HRC might benefit individuals pursuing the creation of

- A genetic replica of a dying child.
- A savior sibling for a sick child affected by some nonhereditary condition.
- Genetically related children (for those who lack a partner, are in a same-sex relationship, or cannot conceive owing to male infertility).
For bioliberals, the same arguments justifying new reproductive technologies in these scenarios also justify HRC, which should be tried when it reaches adequate levels of safety (Pence, 1998).

Bioconservatives, who are pessimistic about the use of genetics and biotechnology on humans, reject that claim: HRC crosses a moral line that should not be crossed. At present, bioconservatism is the prevailing official view in public policy, as can be inferred from widespread bans and moratoria on HRC. The strongest objection to it concerns preventing harm to the clone, which currently would be at high risk. This suggests that researchers should not attempt clinical trials now. But suppose cloning technology develops to a point where the risk becomes low. Then what? Some bioconservatives would still object, on the grounds of

1. People's reactive attitudes toward HRC,
2. Slippery-slope scenarios, and
3. HRC's hubristic motivation.

Leon Kass makes all three objections. Aware that even if pervasive, reactive attitudes such as repugnance depend on intuitions not everyone shares, Kass contends that they are enough to put the burden of argument on bioliberals. Yet since bioliberals can appeal to HRC's benefits and procreative freedom, (1) has limited impact. To succeed, bioconservatives need compelling reasons for (2) and (3). The slippery slopes in (2) consist of futuristic scenarios where the parent-child relationship is severed because of factors such as the asexual mode of conception involved in HRC and the confused family links between clones and cell donors. But pessimistic predictions about psychological harm to children born to, or growing up in, nontraditional families were made before, in connection with complex cases of assisted reproduction. Since such predictions failed to be borne out by the evidence, more is needed to support this slippery slope argument against HRC.

Fueling (3), the Hubristic Motivation charge, are worries about the rationale of using genetics to control reproduction, and ultimately, nature. For HRC, the worries may hinge on its unnaturalness as a means for reproduction. In Chapter 14, a similar objection to the new reproductive technologies was met with the bioliberal reply that it cannot be consistently maintained. After all, unnaturalness also applies to vaccination, fighting epidemics, even taking aspirin. But the worries might be that like PGD, HRC is old eugenics in disguise, or expresses a Promethean aspiration to make reproduction fit our desires. We take up each of these objections to reprogenetics in turn.

**Early Eugenics**

Before the development of contemporary genetic engineering, early eugenicists attempted to manipulate the genome indirectly, by means of controlled breeding and
selection of offspring. Genetic engineering, sometimes called 'modern eugenics,' opened a road for the direct manipulation of an organism's gene (or genes) in order to alter it in a controlled way. The resulting genetically modified organism has or lacks some targeted genetic traits.

Genetic engineering for humans may be either

1. Positive—aimed at enhancing genetic makeup, or
2. Negative—aimed at treating or preventing genetic disorders.

The implementation of (1) and (2) during human reproduction raise moral issues. In particular, positive genetic engineering amounts to a new eugenics in which genetic engineering technology is put at the service of human enhancement. A host of concerns about (1) stem from the catastrophic moral failure of early eugenics programs. In fact, the word 'eugenics' (literally 'good origins') acquired its bad reputation in the early 20th century, when it came to designate reproductive policies mandated by the state for the creation of 'better' humans. The early eugenicists hoped to produce better humans by 'controlled breeding' aimed at selecting individuals with the 'best' genes.

Their programs, imbued with end-of-the-century optimism and pseudo-science, spread rapidly in the US, reaching their peak in the early 1920s. Similar programs were in place in Canada, Denmark, Finland, Sweden, Norway, and Germany. All traded on existing racial prejudices against 'inferior' races and disabled people, many of whom were forcibly prevented from reproducing and often confined to institutions. In the US, after World War I a number of pro-eugenics societies, active in most states, advocated preventing the reproduction of people with 'heritable degenerative' traits such as epilepsy, alcoholism, or poverty. Their agenda influenced laws justifying restrictions on marriage and involuntary sterilization. Indiana enacted a sterilization law in 1907. By 1931, 30 states had eugenic laws and about 12,000 sterilizations had been performed, half of them in California. In 1924, the US Congress passed the Immigration Act, and President Coolidge signed it into law, remarking that "America should be kept American. ..." In his view, a national origins quota was needed for immigrants from groups considered undesirable such as Southern Europeans and Asians.

The National Socialist movement that came to power in Germany in 1933 took the early eugenics program to an extreme. For Nazi doctors and political leaders, eugenics was the tool for achieving a racially homogenous, 'superior,' 'Aryan' society. It required discrimination, exclusion, and ultimately the elimination of those deemed unfit for that society, which included Germans judged 'congenitally' sick, asocial, or homosexual as well as Communists, Jews, and people from other ethnic groups such as the Roma and Slavs of all nationalities. Murdering mentally impaired individuals was part of Nazi eugenics. In a single institution that is now a museum, 40,000 of them were killed.
As the eugenics agenda was evolving into a program of mass extermination in the hands of the Nazis, it was withering in North America under internal criticism, in part for its pseudo-scientific basis and its coercive character. Geneticists now believe that early eugenicists misunderstood Darwin and Mendelian laws of heredity. But Michael Sandel suspects that they never put eugenics to rest. Their eugenicist pursuits continued to operate privately, not by coercive state policies, but embedded in the rationale for implementing genetic technology in order to gain control over implementing genetic technology in order to gain control over reproduction for no medical reasons, such as in some uses of the technology we consider next.

**IMAGE 15.1**
©iStockPhoto/RTImages

15.3 Human Genetic Enhancement

*The Posthuman Scenario*

Positive genetic engineering has already shown promising results in experiments on nonhuman mammals. Studies on mice, for example, have shown that their memory and learning abilities can be greatly enhanced through genetic alterations. Similar alterations might one day be feasible in humans. As with human cloning, the prospect of human genetic enhancement (HGE) is exciting for bioliberals and alarming for bioconservatives. HEG may involve inter-species genetic transfer, the so-called
chimeras produced by hybrid human-animal embryos, or the eugenic selection of traits for offspring by means of PGD and in-vitro fertilization— which may one day permit choosing embryos for nondisease traits such as hair or eye color, complexion, athletic ability, intelligence, and even personality. Gene editing and other technological innovations might make possible the creation of genetic sequences and protein structures associated with powers and immunities no human being has ever had. Individuals so engineered will be posthumans since their physical or mental capacity will be improved to a level beyond what humans can achieve without HGE.

For enthusiasts, it is only a matter of time until genetics makes the posthuman scenario possible. That would, however, require genetic knowledge well beyond that available today, since posthumans will be endowed with features beyond the average for present-day humans, such as being more long-lived, disease-resistant, intelligent, good looking, and emotionally stable, or having enhanced athletic abilities, heightened acuity for aesthetic appreciation, and greater moral discernment.

**Defining 'Enhancement'**

A genetic enhancement need not be so ambitious as to provide powers and immunities no human being has ever had. It may merely provide or augment a trait deemed advantageous or pleasing within a family or larger group. Such modifications count as
Posthuman

Any individual genetically modified to have capacities, powers, or immunities well beyond human standard endowment.

enhancements whenever they aim at improving average capacity for the sake of providing such positional goods. They go beyond genetic therapy—i.e., modifications for treating or preventing disease. Unlike genetic enhancement, therapeutic genetic modifications are often considered morally permissible.

But the alleged moral distinction between genetic enhancement and therapy is affected by vagueness and relativity, as can be seen when applied to genetic modification of human intelligence. According to the IQ scale, average intelligence falls between the 90 and 109 scores. Given the enhancement-versus-treatment distinction, raising someone's IQ score from 69 (mentally impaired) to 109 (normal intelligence) counts as morally permissible treatment, but raising it to 110 (the lower limit of superior intelligence) is morally impermissible enhancement. How could one IQ point make such a big moral difference? Here the problem is vagueness. Now let's consider the relativity problem, which arises because the moral permissibility of raising someone's IQ by, say, 40 points would vary from person to person. If a person's initial IQ score is 69, raising it to 109 is permissible therapy. But if it is 100, raising it to 140 is impermissible enhancement. Therefore, the moral permissibility of augmenting intelligence by exactly 40 IQ points varies from person to person. Proponents of the distinction will need a clear conceptualization of 'average capacity' to handle these problems.

A further complication arises because there might be moral differences among different types of human enhancement. Not all enhancements need be genetic. For example, athletic performance can also be enhanced by training (weight lifting), equipment (e.g., running shoes, special swimming suits), and drugs (doping). Although not all such enhancements face moral objections, genetic ones often do. In addition, Kamm (2005) classifies enhancements as 'ex ante' or 'ex post,' depending on whether they occur before or after the enhanced child exists. Ex ante changes are strictly genetic, while ex post changes also include drugs, training, and equipment. Let's now look closely at the moral controversy involving ex ante enhancements, the so-called designer babies.

**Designer Babies**

Advances in genetic engineering might one day make it possible to create designer babies. The scope of powers and capacities eligible for such genetic enhancements is quite wide. They may involve

- Physical appearance—e.g., eye color, complexion, hair color.
• Artistic abilities—e.g., musicality, appreciation of the visual arts.
• Psychology—e.g., intelligence, optimism, emotional stability.
• Morality—e.g., character, motivation to act morally, moral discernment.
• Health condition—e.g., life span, disease resistance.

Conversely, parents might want genetic alterations for their children to eliminate negative traits such as violence or depression. At present, of course, genetic knowledge enables enhancements of neither type, a situation that is likely to prevail for at least the immediate future. Consider, for instance, intelligence, whose genetic basis is still largely unknown. It is unlikely that intelligence involves a single gene that will soon be discovered by behavioral geneticists, or that it will be amenable to risk-free enhancement. As noted by cognitive psychologist Stephen Pinker, gene editing aimed at enhancing intelligence might involve risks of causing retardation that no parent will want to take. But let’s assume that one day such enhancements become both affordable and medically safe. Should they be an option?

**Bioliberalism about HGE**

Recasting the question as ‘Should we take that opportunity?’ bioliberals answer, ‘Yes.’ Savulescu endorses this answer with a so-called Perfectionist Argument for genetic enhancements contending that the parents have a prima facie obligation to enhance their offspring as much as possible, a conclusion that follows from the above discussed principle of Procreative Beneficence, according to which parents should select, among possible children, the child with the best chance to have the best life (given the available information). To support this argument, Savulescu invokes intuitions about these cases:

**Rubella.** A woman contracts rubella. If she conceives now, the baby’s chances of severe mental impairment, blindness, and deafness are high. If she waits three months, she will conceive a different but healthy child.

**Asthma.** A couple have two IVF embryos, one with no abnormalities, the other showing a predisposition to develop asthma.

Procreative Beneficence captures common intuitions about **Rubella** (viz., that she should wait three months) and **Asthma** (viz., that they select embryo A). Yet perfectionists face some problems, since they are committed to the selection of offspring for trivial reasons. Consider

**Pizza.** A woman might give birth to either child A or child B. A has a slightly higher chance of enjoying pizza compared to child B. Other things are equal.

The intuition here is that there is no parental duty to select child A. But, given the Perfectionist Argument, there is. Thus, perfectionism is not always in agreement with
intuition. Furthermore, its appeal to cases such as Asthma oversimplifies what will be involved in selecting the best child. Genetic knowledge is likely to yield more complex choices than child-with-asthma-versus-healthy-child, since it might reveal that child A will develop an early, intractable form of cancer while child B an also intractable mental illness. If such choices become the norm, the Perfectionist Argument will be silent most of the time. In addition, perfectionism might have ill effects for society, since as discussed in connection with PGD, it might justify selecting children for objectionable reasons such as skin color or sex. It might also produce an unjust social order where enhancement will be accessible only for the wealthy. Perfectionists can meet some of these objections by invoking public policy.

Yet, taken together, the objections suggest that perfectionism is too strong, especially compared with libertarianism, which makes the weaker claim that enhancing children is morally permissible. Consistent with their appeal to Procreative Freedom, libertarians think that parents have a right to choose their children, provided their choice does not harm the children or others. Another less ambitious alternative to perfectionism is sufficientarianism, which holds that the parental duty to enhance need not aim at perfection, but only at reaching a sufficient level whereby the child could achieve sufficient positional goods. In the sufficientarian view, ‘best’ becomes ‘decent,’ so that parents have the prima facie obligation to select the child who has a decent chance of having a sufficiently decent life. Sufficientarianism accommodates the intuitions that in Pizza, the parents have no duty to select child A, while in Rubella they have the duty to wait. What the view would say about Asthma depends on what a ‘sufficiently decent’ life is, something that needs clarification.

Modest Bioconservatism about HGE

Bioconservatives can be modest or extreme, depending on whether they prescribe only caution about HGEs or forbid it altogether. Modest bioconservatives see nothing intrinsically wrong with HGE but they accept this general rationale for the genetic engineering of any organism:

**Precautionary Principle**

No positive (nontherapeutic) genetic engineering should be attempted without sound evidence that its potential benefits can outweigh bad outcomes.

Given this principle, researchers should simply steer clear of attempting any HGE until there is solid evidence of its safety. Current gene therapy on humans illustrates the potential risk of gene manipulation, about which we'll have more to say later. And as noted earlier, gene editing to enhance intelligence might carry the risk of causing retardation. More generally, genetic modifications that alter the course of natural selection might destroy the human race.
Bioliberals reject such 'alarmist' predictions, charging that the Precautionary Principle has paralyzing effects. If followed consistently, it would result in the starvation of the developing world. For it will not allow for genetically modified crops, which are the only hope for meeting the demands of an overpopulated planet. However, by distinguishing levels of risk, modest bioconservatives can hold some genetic modifications permissible while also allowing that HGE is still too risky as a general policy, and therefore morally forbidden. Modest bioconservatives also reject the extreme bioconservative claim that HEG is wrong per se.

**Extreme Bioconservatism**

These bioconservative arguments aim at substantiating HGE’s intrinsic wrongness. For the Social Justice argument, were HGE to become an accepted practice, it would deepen economic and social injustice. It may even lead to a bifurcation of the human race into two sub-species, the enhanced posthumans (descendants of those who could afford to get enhanced) and the unenhanced underclass (everybody else). In reply, bioliberals note that this frightening, Brave-New-World scenario is just one of the possible outcomes. After all, it cannot be ruled out that public policies might be in place to ensure more egalitarian distribution of HGE, or a prioritarian distribution according to need. Besides, we already have, in the industrial democracies of the developed world, a two-tiered social-class system (much of it rooted in genetics, owing to people’s preference to marry within their social class) where most benefits and burdens are distributed unevenly along class lines.

The Eugenics objection to HGE charges that it is an enterprise rooted largely in normative judgments, a slippery foundation if ever there was one. Different people will have different views about what traits should be selected for enhancement, reduction, or elimination. Considering the historical abuses of early eugenics reviewed above, who should determine which traits fall in what category? Furthermore, we may set out to change human nature for the better, but are we really confident that we know what we should change it to? Suppose everybody will agree that a tendency toward unprovoked, aggressive violence is not a desirable human trait, for instance, or that an inclination toward cooperation is desirable. Even so, there will also be countless human traits held desirable by some and condemned by others. How tall, aggressive, tolerant, or risk-averse should one be?

On the other hand, some bioconservatives predict a Human Diversity problem facing HGE. According to this objection, the aggregate effects of individual choices will be bad because they will diminish human diversity or eliminate it altogether. Most parents will select similar enhancements for their offspring, something supported by evidence from egg and sperm-donor sales. As a result, most children will have similar traits: blond, tall, etc.

The Eugenics objection and the Human Diversity problem are slippery-slope
arguments making empirical predictions that could turn out to be false. We need to wait and see. Their predictions might never materialize or be correctible by public policy. But other extreme bioconservatives do not make such predictions. For example, Michael Sandel's "deepest objection" focuses entirely on the wrongness of designer babies, and more generally of HGE, given the underlying hubristic attitude HGE expresses. According to this Hubristic Motivation Charge, its wrongness

...lies less in the perfection it seeks than the human disposition it expresses and promotes. The problem is not that parents usurp the autonomy of a child they design. The problem is in the hubris of the designing parents, in their drive to master the mystery of birth ... it would disfigure the relation between parent and child, and deprive the parent of the humility and enlarged human sympathies that an openness to the unbidden can cultivate. (Sandel, 2004: 57)

Enhancements made to satisfy the parents' own aesthetic or social preferences about their children's eye color or athletic ability seem wrong on a number of counts. For one thing, these are shallow values, and even outright social prejudices. Furthermore, such enhancements may treat children as mere instruments for their parents' enjoyment and thus commodify them. And by eliminating the element of chance in the natural lottery, they may also undermine the parent-child relationship.

To these criticisms, Sandel adds HGE's underlying hubristic motivation. It expresses a disposition toward mastering or controlling reproduction, one of the mysteries of nature with which we should not interfere. For him, mastery stands in contrast to "openness to the unbidden," which requires acceptance of nature's gifts and provides the basis for societal values such as humility, responsibility, and solidarity. Openness to the unbidden requires a balance between the "acceptance" and the "transforming" love parents owe to their children. One amounts to unconditional love, the other allows for seeking some enhancements short of genetic ones. Permissible enhancements are achieved through education, training, and equipment. For example, enhancing speed through the use of proper running shoes is morally permissible, but designing a baby with fast-twitch muscle fiber is not. Both actions enhance speed, a natural gift, and may result in exactly the same speed increase. What makes one permissible and the other not? Maybe that HGEs are more permanent. But this cannot be so, since many offspring alterations short of genetic enhancement also result in permanent, heritable changes. Environmental factors such as education and training can cause permanent alterations in a person's neurophysiology that will be passed on to succeeding generations. Since Sandel considers these transformations permissible, consistency requires ruling out the permanence of HGE as the reason for its impermissibility. But Sandel provides no principle for consistently distinguishing between permissible and impermissible enhancements.

Furthermore, note that agents' motivation may be hubristic or otherwise bad
without this implying that their actions thus motivated are wrong. In the early 1900s, at the end of the Spanish-American War, Walter Reed and his medical staff performed experiments in Cuba that established yellow fever as a mosquito-borne illness. They exposed themselves and others to mosquito bites to test this hypothesis. Deemed a scientific success, their experiments prompted mosquito-control measures that greatly reduced the incidence of yellow fever. Traditionally, the team’s motivation was associated with honor and bravery. But historians of medicine now believe that it included a great deal of professional and occupational self-interest, together with monetary compensation. Plainly, this assessment of the investigators’ attitude changes how we should evaluate their moral character, but not the value of their experiments to potential victims of yellow fever, and thus, to society.

15.4 Chapter Summary

This chapter first looked at arguments for and against the moral permissibility of parental decisions to use preimplantation genetic diagnosis (PGD) to select for the presence of certain disability traits in IVF embryos. The Autonomy Argument finds these contrary to a child’s right to an open future. Reasons for the moral permissibility of such decisions invoke procreative freedom and genetic essentialism. In the deaf couple case, some libertarians and utilitarians contend that a different parental procreative decision would have produced either no child or another child. Since the deaf child’s life is worth living, the couple’s choice was permissible. A sticking point for utilitarians is whether all harms are person-affecting. If not, then designing disabled children is wrong when it brings about a less-than-optimal state of affairs.

For PGD critics, its use in selecting embryos for nonmedical traits such as sex (in the absence of risk for sex-related genetic disease) expresses a hubristic attitude analogous to that expressed by early eugenicists. This attitude fuels also human reproductive cloning (HRC) and human genetic enhancement (HGE). HRC defenders argue that it’ll be just another new reproductive technology, justifiable by its potential benefits. Critics invoke public reaction to HRC, concerns about its potential harm (to clones, families, or society), and/or the underlying hubristic disposition toward controlling nature that HRC expresses. Since each of these objections can be met but the evidence from animal cloning suggests HRC is presently unsafe, experimentation (if any) should proceed only with proper oversight.

Like HRC, HGE involves speculative scenarios which might one day become real. While bioconservatives consider impermissible the enhancement of physical and mental capacities, for bioliberals it is permissible and even mandatory. In both camps, there are extreme and modest positions, with only the modest ones surviving criticism. In fact modest bioliberalism (sufficienarianism) can be combined with modest conservatism (the precautionary principle) to contend that HGEs aiming at improvement, not perfection, are morally permissible but should not be administered
until they become safe. Here is a roadmap of these positions.

**FIGURE 15.2** Main Arguments for and against the Moral Permissibility of HGE
The Basic Argument for Vegetarianism
James Rachels identifies Peter Singer’s argument that it is wrong to cause pain unless there is a good enough reason as the argument which convinced him to become a vegetarian. Rachels describes some of the facts involved in the meat-production industry and maintains that such facts are a vital part of the argument. Since the facts are well established, how can we account for so many people being unmoved by these facts? According to Rachels, there is a difference between how scientists and animal-rights advocates think about the cognitive abilities of nonhumans. But for Rachels the important issue is animal pain.

In 1973 Peter Singer, who was then a young, little-known philosopher from Australia, published an article called “Animal Liberation” in the New York Review of Books. The title suggested that there was a parallel between our treatment of animals and the unjust treatment of blacks and women. At first, it was hard to take the comparison seriously. Many proponents of “black liberation” and “women’s liberation,” as those movements were then known, found the comparison insulting, and most philosophers thought the topic was hardly worth discussing. But Singer kept at it, writing more articles and a now-famous book. It is now commonly said that the modern animal-rights movement grew out of those works. Thanks to Singer, many people, including me, became convinced that a fundamental change in our attitude toward animals was necessary. The indispensable first step was becoming a vegetarian.

The argument that persuaded me to become a vegetarian was so simple that it needs only a little elaboration. It begins with the principle that it is wrong to cause pain unless there is a good enough reason. The qualification is important, because causing pain is not always wrong. My dentist causes me pain, but there’s a good reason for it, and besides, I consent. My children’s doctor caused them pain when he gave them their shots, and they did not consent, but that was all right, too. However, as the principle says, causing pain is acceptable only when there is a good enough reason for it. Justification is required.

The second step in the argument is to notice that in the modern meat-production business, animals are made to suffer terribly. There is a reason for this suffering, too. We eat the meat, and it helps to nourish us. But there is a catch: we could just as easily nourish ourselves in other ways. Vegetarian meals are also good. Nonetheless, most people prefer a diet that includes meat because they like the way it tastes. The question, then, is whether our enjoyment of the way meat tastes is a good enough reason to justify the amount of suffering that the animals are made to endure. It seems obvious that it is not. Therefore, we should stop eating the products of this business. We should be vegetarians instead.

I will call this the basic argument. It has a limited application. It says nothing
about animals raised on old-fashioned family farms or animals killed in hunter-gatherer societies. It addresses only the situation of people like us, in modern industrial countries. But it does point out, in a simple and compelling way, why those of us in the industrial countries should not support the meat-production business as it now exists.

When I emphasize the argument's simplicity, I mean that it does not depend on any controversial claims about health or on any religiously tinged notions of the value of life. Nor does it invoke any disputable ideas about "rights." Further claims of these kinds might strengthen the case for vegetarianism, but the basic argument does not depend on them. Nor does it rest on any contentious philosophical theory about the nature of morality. Philosophers sometimes misunderstand this when they think it is a merely utilitarian argument and that it can be refuted by refuting utilitarianism. But the basic argument is not tied to any particular theory about the nature of ethics. Instead, it appeals to a simple principle that every decent person already accepts, regardless of his or her stand on other issues. The most striking thing about the argument is that it derives such a remarkable conclusion from such a sober, conservative starting point.

The basic argument, then, is common ground for people of various moral and political persuasions. Matthew Scully is in most respects the antithesis of Peter Singer. Scully, a former speechwriter for various Republicans including President George W. Bush, recently surprised his conservative friends by writing a book, *Dominion: The Power of Man, the Suffering of Animals, and the Call to Mercy*, in which he detailed the cruelties of the modern factory farm—cruelties that are, in his words, "hard to contemplate." Scully reports:

Four companies now produce 81 percent of cows brought to market, 73 percent of sheep, half our chickens, and some 60 percent of hogs. From these latter, the 355,000 pigs slaughtered every day in America, even the smallest of mercies have been withdrawn. In 1967 there were more than a million hog farms in the country; today there are about 114,000, all of them producing more, more, more to meet market demand. About 80 million of the 95 million hogs slaughtered each year in America, according to the National Pork Producers Council, are intensively reared in mass-confine ment farms, never once in their time on earth feeling soil or sunshine. Genetically engineered by machines, inseminated by machines, monitored, herded, electrocuted, stabbed, cleaned, cut, and packaged by machines—they themselves treated as machines "from birth to bacon"—these creatures, when eaten, have hardly ever been touched by human hands.

Scully visited some of these automated pig farms in North Carolina, and his report is chilling. Sows have been engineered to weigh 500 pounds each. Pigs are crowded twenty each in pens only seven-and-a-half feet square. The close confinement creates problems in managing the animals. Pigs are intelligent and social animals who
normally build nests and keep them clean. They will not urinate or defecate in their nests, as they must do in the pens. They form bonds with other animals. They want to suck and chew, but in the pens, being deprived of a normal environment in which they can do these things, they begin to chew on the tails of the animals in front of them. In such close quarters, the victims cannot escape. The chewing causes infection, and sick pigs are no good. The solution is “tail docking,” a procedure recommended by the U.S. Food and Drug Administration, in which the pigs’ tails are snipped (without anesthetic) by pliers. The point is to make the tails more sensitive to pain, so that the animals will make a greater effort to avoid their neighbors’ attacks. Surveying the whole setup, the operator of one such “farm” observes: “It’s science driven. We’re not raising pets.”

When critics of the meat-production industry report such facts, their accounts are often dismissed as “emotional appeals.” But that is a mistake. It may be true that such descriptions engage our emotions. However, emotionalism is not the point. The point is to fill in the details of the basic argument. The basic argument says that causing pain is not justified unless there is a sufficiently good reason for it. In order to apply this principle to the case of factory farming, we need to know how much pain is involved. If only a little pain were being caused, a fairly insubstantial reason (such as our gustatory pleasure) might be sufficient. But if there is extensive suffering, that reason is not enough. Thus, these facts are a vital part of the argument, and it is necessary to keep them in mind when considering whether the argument is sound. For those of us who have no firsthand knowledge of the subject, reports by such relatively impartial observers as Matthew Scully are indispensable. Another report recently appeared in the New York Times Magazine. The author, Michael Pollan, went to a great deal of trouble to find out what happens to cattle who are raised and slaughtered for beef. “Forgetting, or willed ignorance, is the preferred strategy of many beef-eaters,” he says; but Pollan wanted to see for himself the conditions in which the animals live and die. So he bought a steer—“No. 534”—at the Blair Brothers Ranch in South Dakota and followed its progress to the slaughterhouse. No. 534 spent the first six months of his life in pastures alongside his mother. Then, having been weaned and castrated, he was shipped to Poky Feeders, a feedlot operation in Garden City, Kansas.

“A cattle feedlot,” says Pollan, “is a kind of city, populated by as many as 100,000 animals. It is very much a premodern city, however—crowded, filthy and stinking, with open sewers, unpaved roads and choking air.” Fecal dust floats in the air, causing irritation to the eyes and lungs. Searching for No. 534, Pollan found his animal standing in a “deep pile of manure.” Dried manure caked on the animals is a problem later, in the slaughterhouse, where steps must be taken to ensure that the meat does not become contaminated. In the feedlot itself, disease would kill the animals were it not for massive doses of antibiotics.

At the Blair Brothers Ranch, No. 534 ate grass and was given corn and alfalfa
hay to fatten him up. In his last six weeks at the ranch, he put on 148 pounds. After being shipped to Poky Feeders, he would never eat grass again. His diet would be mostly corn and protein supplement, "a sticky brown goop consisting of molasses and urea." Corn is cheap, and it produces "marbled" beef, although it is not what the animals naturally desire. In a grisly sort of forced cannibalism, the animals are also fed rendered cow parts. The animals could not live on this diet for long—it would "blow out their livers," said one of the feedlot operators. But they are slaughtered before this can happen. The diet is effective, however: the animals weigh more than 1,200 pounds when taken to the slaughterhouse.

No. 534 was slaughtered at the National Beef Plant in Liberal, Kansas, a hundred miles down the road from Poky Feeders. This is where Pollan’s personal observations come to a stop. He was not allowed to watch the stunning, bleeding, and evisceration process; nor was he permitted to take pictures or talk to the employees.

Opposing cruelty should not be seen as a specifically liberal or conservative cause. Scully, the conservative Republican, emphasizes that one should oppose it "even if one does not accept [the animal rights advocates'] whole vision of the world." He makes a point of distancing himself from Peter Singer, who champions various left-wing causes. Singer is wrong about the other issues, says Scully, but he is right about the animals.

II

The basic argument seems to me obviously correct. But its very obviousness suggests a problem: if it is so simple and obvious, why doesn’t everyone accept it? Why doesn’t everyone who has this argument explained to them become a vegetarian? Of course, many people do, but most do not. Part of the explanation may be that it is natural for people to resist arguments that require them to do things they don’t want to do. If you want to go on eating meat, you may pay no attention to arguments that say otherwise. Moreover, people generally do not respond to ethical appeals unless they see others around them also responding. If all your friends are eating meat, you are unlikely to be moved by a mere argument. It is like an appeal for money to provide vaccinations for third-world children. The argument that the vaccinations are more important than your going to a movie may be irrefutable, considered just as an argument. But when no one around you is contributing, and your friends are all going to the movie, you are likely to ignore the charitable appeal and spend the money on popcorn instead. It is easy to put the children out of mind.

All this may be true. But there is a more pressing problem about the basic argument—at least, a more pressing problem for me, as a philosopher. Many of my professional colleagues are unmoved by this argument, and I am not sure why. Those who study ethics, especially from a nonreligious point of view, often find the argument compelling. But others do not. This is puzzling because professional
philosophers—those who teach in colleges and universities—study arguments dispassionately, and while they often disagree, they disagree about arguments only when the issues are tricky or obscure. But there is nothing tricky or obscure about the basic argument. Thus I would expect that, on so simple a matter, there would be widespread agreement. Instead, many philosophers shrug the argument off.

The same is true of other academics who study cognitive science, psychology, and biology. They are at least as smart as I am, if not smarter, and they are morally decent people. Yet, while I think the basic argument is compelling, many of them do not. It is not that they think the argument makes a good point, even though they are unwilling to act on it. Rather, they find the argument itself unconvincing. How can this be?

Sometimes philosophers explain that the argument is unconvincing because it contains a logical gap. We are all opposed to cruelty, they say, but it does not follow that we must become vegetarians. It only follows that we should favor less cruel methods of meat production. This objection is so feeble that it is hard to believe it explains resistance to the basic argument. It is true enough that if you are opposed to cruelty, you should prefer that the meat-production business be made less brutal. But it is also true that if you are opposed to cruelty, you have reason not to participate in social practices that are brutal as they stand. As it stands, meat producers and consumers cooperate to maintain the unnecessary system of pig farms, feedlots, and slaughterhouses. Anyone who finds this system objectionable has reason not to help keep it going. The point would be quickly conceded if the victims were people. If a product—curtains, let’s say—were being produced by a process that involved torturing humans, no one would dream of saying: “Of course I oppose using those methods, but that’s no reason not to buy the product. After all, the curtains are very nice.”

Many in the animal-rights movement believe that scientists are blinded by the need to justify their own practices. The scientists are personally committed to animal experimentation. Their careers, or the careers of their colleagues, are based on it, and they would have to stop this research if they conceded that animals have moral claims on us. Naturally they do not want to do this. Thus they are so biased in favor of current practices that they cannot see the evil in them. This explains why they cannot see the truth even in something so simple as the basic argument.

Perhaps there is something to this, but I do not want to pursue it. On the whole it is a condescending explanation that insults the scientists, cuts off communication with them, and prevents us from learning what they have to teach us. It should be noted, however, that the basic argument about vegetarianism is independent of any arguments about animal experimentation. Indeed, the case against meat eating is much stronger than the case against the use of animals in research. The researchers can at least point out that in many instances, their work has a serious purpose that can
benefit humankind. Nothing comparable can be said in defense of meat eating. Thus, even if some research using animals were justified, meat eating would still be wrong.

I believe a better explanation is in terms of the overall difference between how scientists and animal-rights advocates think about the nature of nonhumans. Defenders of animal rights tend to see the differences between humans and nonhumans as slight. They frequently emphasize how much the animals are like us, in order to argue that our ethical responsibilities to the animals are similar to our responsibilities to one another. Animals are pictured as intelligent and sociable creatures who love their children, who experience fear and delight, who sulk, play, mourn their dead, and much more. So how can it be denied that they have rights, just as we do? I have argued in this way myself, more than once.

Many scientists, however, see this as naive. They believe the differences between humans and other animals are vast—so vast, in fact, that putting humans in a separate moral category is entirely justified. Moreover, they feel they have some authority on this score. After all, the scientific study of animals is their professional concern. In light of this, how should we expect them to react when they are confronted by belligerent amateurs who insist they know better? It is only natural that the scientists should disregard the amateurs' arguments.

A case in point is the anthropologist Jonathan Marks, who teaches at the University of North Carolina at Charlotte. In 1993, Peter Singer and Paola Cavalieri, an Italian writer on animal issues, initiated a campaign known as "the Great Ape Project," an effort to secure basic rights for our closest relatives, the chimpanzees, gorillas, and orangutans. The rights being demanded were life, liberty, and freedom from torture. Marks was invited to participate in a debate about these demands, and he recorded his thoughts in an engaging book, What It Means to Be 98% Chimpanzee. "Since their brains are closely related to our brains," Marks says, "it should come as no surprise that the apes can approach humans in their cognitive functions." Despite this, "Apes are often objectified by callous and cynical entrepreneurs, who neither regard them nor treat them as the sentient, emotionally complex creatures they are." Marks does not think this is acceptable. "Apes deserve protection," he says, "even rights."

Reading these words, one would expect Marks to be an ally of Singer and Cavalieri. But he is not. The Great Ape Project, he thinks, is completely wrongheaded. Why? Marks's attempt at philosophical argument is unimpressive—he says the critical issues are that chimps, gorillas, and orangutans aren't human, and that in any case we are politically powerless to guarantee such rights even for humans. Of course, these arguments get us nowhere. Everyone knows the animals aren't human; the point is that they are sufficiently like humans to deserve the same basic protections. And the fact that we cannot ensure rights for humans does not mean that we should stop thinking humans ought to have them.
The underlying reason for Marks’s scorn of the animal-rights ideology becomes clear when he turns to the scientific study of animal behavior. The similarities between humans and other great apes, he intimates, are only superficial: “Where clever, controlled experimentation has been possible, it has tended strongly to show that in specific ways, ape minds work quite differently from human minds.” For support, he cites the work of the psychologist Daniel J. Povinelli, who argues that chimpanzees’ conceptions of physical interactions (as, for example, when a hook is used to manipulate an object) are very different from human understanding. Marks does not say how this fits with his earlier assertion that “apes deserve protection, even rights,” but clearly, in his view, the latter thought trumps the former.

We find this pattern repeated again and again: The scientists concede that the animal rights advocates have a bit of a point, but then the scientists want to talk about the facts. They think we do not know nearly enough about the details of how animal minds work to justify any firm moral conclusions. Moreover, such knowledge as we do have suggests caution: the animals are more different from us than it seems. The advocates of animal rights, on the other hand, think the facts are well enough established that we can proceed without further ado to the ethical conclusions. Anyone who suggests otherwise is viewed as dragging his/her feet, perhaps to avoid the unpleasant truth about the injustice of our behavior toward the animals.

III

What are we to make of all this? One obvious idea is that we should take seriously what the scientists tell us about what animals are like and adjust our moral conceptions accordingly. This would be an ongoing project. It would take volumes even to begin, by considering what is currently known. But those volumes would be out of date by the time they were completed, because new discoveries are being made all the time.

However, where the basic argument is concerned, the only relevant part of this project would be what science can tell us about the capacity of animals to experience pain. Jeremy Bentham famously said, “The question is not, Can they reason? nor Can they talk but Can they suffer?” To this we might add that, contrary to Jonathan Marks, it is irrelevant whether chimps have a different understanding of physical interactions. It is irrelevant, that is, if we are considering whether it is acceptable to treat them in ways that cause them pain.

This point is easily misunderstood, so it is worth elaborating just a bit. Of course, the facts about an individual are important in determining how that individual should be treated. (This is true of humans as well as nonhumans.) How an animal should be treated depends on what the animal is like—its nature, its abilities, and its needs. Different creatures have different characteristics, and these must be taken into account when we frame our ethical conceptions. The scientific study of animals gives
us the factual information we need. But not every fact about an individual is relevant to every form of treatment. What facts are relevant depends on what sorts of treatment we are considering. To take a simple example, whether an animal can read is relevant if we are considering whether to admit him to university classes. But the ability to read is irrelevant in deciding whether it is wrong to operate on the animal without anesthesia. Thus, if we are considering whether it is wrong to treat pigs and cattle in the ways we have described, the critical issue is not whether their minds work in various sophisticated ways. The critical issue is, as Benthem said, whether they can suffer.

What does science tell us about this? The mechanisms that enable us to feel pain are not fully understood, but we do know a good bit about them. In humans, nociceptors—neurons specialized for sensing noxious stimuli—are connected to a central nervous system, and the resulting signals are processed in the brain. Until recently it was believed that the brain’s work was divided into two distinct parts: a sensory system operating in the somatosensory cortex, resulting in our conscious experiences of pain, and an affective-motivational system associated with the frontal lobes, responsible for our behavioral reactions. Now, however, this picture has been called into question, and it may be that the best we can say is that the brain’s system for processing the information from the nociceptors seems to be spread over multiple regions. At any rate, the human nociceptive system also includes endogenous opioids, or endorphins, which provide the brain with its natural painkilling ability.

The question of which other animals feel pain is a real and important issue, not to be settled by appeals to “common sense.” Only a completed scientific understanding of pain, which we do not yet have, could tell us all that we need to know. In the meantime, however, we do have a rough idea of what to look for. If we want to know whether it is reasonable to believe that a particular kind of animal is capable of feeling pain, we may ask: Are there nociceptors present? Are they connected to a central nervous system? What happens in that nervous system to the signals from the nociceptors? And are there endogenous opioids? In our present state of understanding, this sort of information, together with the obvious behavioral signs of distress, is the best evidence we can have that an animal is capable of feeling pain.

Relying on such evidence, some writers, such as Gary Varner, have tentatively suggested that the line between animals who feel pain and those who do not is (approximately) the line between vertebrates and invertebrates. However, research constantly moves forward, and the tendency of research is to extend the number of animals that might be able to suffer, not decrease it. Nociception appears to be one of the most primitive animal systems. Nociceptors have now been identified in a remarkable number of species, including leeches and snails.

The presence of a perceptual system does not, however, settle the question of whether the organism has conscious experiences connected with its operation. We
know, for example, that humans have perceptual systems that do not involve conscious experience. Recent research has shown that the human vomeronasal system, which works through receptors in the nose, responds to pheromones and affects behavior even though the person is unaware of it. (It was long believed that this system was vestigial in humans, but it turns out that it is still working.) The receptors for “vomerolfaction” are in the nostrils, alongside the receptors for the sense of smell; yet the operation of one is accompanied by conscious experience, while the operation of the other is not. We do not know why this is so. But this suggests at least the possibility that in some species there may be nociceptive systems that do not involve conscious experiences. In that case, those animals might not actually feel pain, even though various indications are present. Is this true of leeches and snails? of snakes? of hummingbirds? We may have strong hunches, but we don’t really know.

Clearly, then, we still have a great deal to learn about the phenomenon of pain in the animal world, and the scientists who work in this area are right to caution us against quick-and-easy opinions. The ongoing study of animal pain is a fascinating subject in itself, and it has enormous importance for ethics. But should this make us less confident of the basic argument? If the issue were our treatment of snails and leeches, perhaps it should. But pigs and cattle are another matter. There is every reason to believe they feel pain—the facts about their nervous systems, their brains, their behavior, and their evolutionary kinship to human beings all point to the same conclusion as common sense: our treatment of them on factory farms and in the slaughterhouses is one of the world’s great causes of misery. If further investigation were to prove otherwise, it would be one of the most astonishing discoveries in the history of science.

Strict vegetarians may want more than the basic argument can provide, because the basic argument does not support sweeping prohibitions. If opposition to cruelty is our motive, we will have to consider the things we eat one at a time. Of course we should not eat beef and pork produced in the ways I have described, and we ought also to avoid factory-farm poultry, eggs, and milk. But free-range eggs and humanely produced milk are all right. Eating shrimp may also turn out to be acceptable. Moreover, from this point of view, not all vegetarian issues are equally pressing: eating fish may be questionable, but it is not nearly as bad as eating beef. This means that becoming a vegetarian need not be regarded as an all-or-nothing proposition. From a practical standpoint, it makes sense to focus first on the things that cause the most misery. As Matthew Scully says, whatever one’s “whole vision of the world” may be, the pig farms, feedlots, and slaughterhouses are unacceptable.
CHAPTER 7

Solidarity Near the End of Life
Summary

Social perspectives on ethical obligations for medical treatment and the care of persons with critical and terminal illness have changed significantly in the past forty years and are in a process of transformation again today. This chapter explores the role of the concept of solidarity and closely related notions in the ethics of end-of-life care. It argues that solidarity is not only relevant to issues of health policy and access to health care but also significant in more personal and intimate settings of clinical and family decision-making. The importance of solidarity has become more apparent as ethical analysis of the goals and practices of end-of-life care have shifted in recent years. This shift has many facets, but in this chapter I characterize it as the transition from an individualistic emphasis on the rights of the dying person to a relational and communicative emphasis on shared decision-making, institutionally shaped transactions (curated caring), and dialogic practices of judgment and decision-making.

Introduction

Social perspectives on ethical obligations for medical treatment and the care of persons with critical and terminal illness have changed significantly in the past forty years and are in a process of transformation again today. This chapter explores the role of the concept of solidarity and closely related notions in the ethics of end-of-life care. It argues that solidarity is not only relevant to issues of health policy and access to health care but also significant in more personal and intimate settings of clinical and family decision-making.

The importance of solidarity has become more apparent as ethical analysis of the goals and practices of end-of-life care has shifted in recent years. This shift has many facets, but in this chapter I characterize it as the transition from an individualistic emphasis on the rights of the dying person to a relational and communicative emphasis on shared decision-making, institutionally shaped transactions (curated caring), and dialogic practices of judgment and decision-making.

As a background to considering this shift in the ethics of end-of-life care, it is important to bear in mind that much work in bioethics has grown out of the liberal Enlightenment tradition in which the freedom and flourishing of the human individual are the primary social and ethical values. Social relationships exist to serve the rational interests of those who are parties to them; individuals do not exist to serve institutions and relationships. This intellectual and philosophical tradition helps explain the appeal of individualistic autonomy and rights to control over what is done by others to one's own body. Such rights have loomed large in end-of-life care, especially once medicine
and technology advanced to a point where these matters could be seriously and regularly at issue. Alternative perspectives that question this ethical individualism are often considered collectivist, communitarian, or metaphysically holistic. The concept of solidarity can easily be classified on the collectivist side of this debate, but in this chapter I hope to show that it would be a mistake to do so.

Neither in theory nor in practice is the emerging relational reorientation in end-of-life medical decision-making a rejection of patient autonomy and the various ethical standards and practical tools that have been developed to hear and heed the voice of those who are dying. But it is a reinterpretation and a reconceptualization of autonomy and the values of independence and subjective choice that have structured the social dynamics of end-of-life care and interpersonal interactions at the bedside. An emphasis on rights and an emphasis on relationships are conceptually distinct, but substantively they point toward many of the same ends. Both seek to foster a social ecology of good dying—by which I mean a relational system of communication, medical decision-making, and treatment that facilitates both the values and obligations of respect for persons and the obligations of care—human presence, relief of suffering, and preservation of dignity. More specifically, the concept of autonomy is central to both rights-based and relational approaches to end-of-life care, but for the former it is an individualistic interpretation of autonomy focusing on noninterference and control, and for the latter it is a relational interpretation of autonomy focusing on contextualized, dialogic judgment and self-affirmation in and through mutuality.

The discussion in this chapter will proceed in three steps. First, I introduce and discuss the concept of solidarity and the ways in which it changes the orientation of ethical analysis relative to the individualistic orientation in bioethics. Next, I explicate in more detail the rise of patient autonomy, patient empowerment, and the right to refuse treatment. Finally, I review and assess the new emphasis on relationality and communication now emerging in end-of-life care. I argue that appeal to the concept of solidarity and autonomy informed by relational judgment can extend and strengthen this new orientation.

My notion is that the challenge facing the relational orientation in end-of-life care is twofold: (1) to provide a more practical and feasible template for medical planning and decision-making that more effectively serves the needs and rights of dying persons; and (2) to avoid bluntly rejecting or abandoning ideals of personal autonomy and respect for persons. The move from rights to relationships can accomplish both of these objectives; it need not purchase the good of practicality or greater patient well-being at the expense of respect for patient agency and personhood. The concept of solidarity can play an important role in meeting both of these objectives. This is due, in part, to the fact that solidarity is constituted by a broader conceptual framework consisting of what I shall call “right recognition” of
moral standing and membership, and "right relationality" of mutual care and concern for others. It is also due to the ways in which solidaristic practices can inform judgment and reflective communication in end-of-life care settings.

**Solidarity: Moral Standing and Standing Up**

Solidarity is a multifaceted concept that is grounded on right recognition and right relationship. Right recognition is a condition of moral membership and equality—namely, the recognition of what has been called the "moral considerability"—the equal and impartial moral standing of each person and the corresponding respect due to the person. Right relationship is a condition of mutuality—the mutuality of interdependence, care for (and of ) others, and concern for their human flourishing, understood in developmental and interdependent terms. It is from these socially and culturally constituted conditions of membership and mutuality that solidarity arises. And it is to them that solidarity makes its basic contribution to ethical theory generally and to the ethics of end-of-life care.

The idea of solidarity carries with it a number of conceptual orientations and commitments. When it interacts with other concepts, such as justice, autonomy, liberty, and dignity, it opens up a relational interpretation of them. For this reason, it is best not to view solidarity as one ethical principle among others. It is not a rule prescribing certain types of conduct or the mental motivation or cause of certain kinds of consequential actions. Nor is it moral virtue in the sense of being a disposition or characterological motivation in the mind of an individual. Solidarity is a social practice. To grasp the role that solidarity can play in the ethics of end-of-life care, it is better to think of ethics discursively rather than architectonically.

Here is what I mean by that distinction. Discursive ethics is a way of articulating and interpreting ethical arguments, vocabularies, standpoints, and narratives; it is a mode of critical and interpretive "discursive practice" that takes place in a historical and social context made up of a number of other contrasting and competing discourses. By a "practice" I mean a normatively structured form of activity (including linguistic and symbolic activity) through which agents pursue a variety of capabilities and activities (or "functionings") that conduce to their human self-realization and constitute a life well lived. Architectonic ethics, by contrast, does not see itself as one discourse among many but instead as a more detached, foundational, and contemplative enterprise. It is concerned with the construction of first principles designed to form the structure and scaffolding for many other, more specific rules. These first principles and secondary rules direct and govern the functioning of institutions and the modes of action and interaction in a society. John Rawls's great work *A Theory of Justice* is an example of architectonic ethics in that he derives two complex principles of justice (on equal liberty and on equal opportunity and distribution of important resources) to oversee
what he calls the "basic structure" of a society. Rawls argues that these principles of justice would emerge from an ahistorical, generic situation of impartial, rational choice and, for that reason, should morally be embraced by members of actually existing societies. (In later work, he argues that these principles reflect and clarify the consensual core of values that contemporary liberal societies have already embraced.)

The discourse and practice of solidarity, on the other hand, do not construct an idealized situation of original choice. It begins with the particularities of society, culture, and psychology—its starting point is the latent possibilities of a given place at a given time and with ongoing forms of conviction and contentment, we might say. It builds on senses of historical memory and tradition, and it feeds on the gratitude felt when one recognizes the service and contributions that others have made to one's way of life. Solidarity begins with the recognition of reciprocal and symbiotic interdependence and then intervenes in—interrupts—an ongoing form of life in order to redress the injustice it contains and win through to a better moral community.

Solidarity tends to move our ethical attention away from the generic, abstract, and universal and toward the practical and concrete particularity of a social, historical, or personal situation. The logic of its conceptual structure is inseparable from what might be called the phenomenological understanding of its practice and the experience of standing in solidarity with others. I shall describe aspects of that phenomenology in a moment. Solidarity inherently leads us to view our own lives and agency as bound together with the rights, well-being, health, and dignity of others here and now. That does not mean that discursive ethics is therefore ethically relativistic, situational, opportunistic, or ad hoc. It means that discursive ethics hazards the wager that the capacity for critical reasoning and moral imagination can grow developmentally out of an engagement with the particularity of the world. However, when you think about it, this is no bolder a wager than the one made by architectonic theorists in their belief that objective moral knowledge can achieve an Archimedean point of eternity and prescind from concrete lived experience in a cultural place or a historical time.

We can begin to get a better sense of what solidarity means by examining what is missing from social and moral life when the idea and the practice of solidarity are absent. Imagine a society in which the well-being of each individual (including one's health status and the circumstances of one's own dying) is seen as sui generis and fundamentally distinct from the well-being of others. Sometimes, to be sure, one person's success may depend in some way on the success and cooperation of others, but then the success of others is valuable to the individual only incidentally. It is as if each person says to herself, if I can do without your cooperation, I shall: all the more to be secure, self-reliant, and independent. There is no moral encumbrance here, no normative push of commitment or obligation, no motivational pull of mutual
recognition and resemblance.

But why then should healthy citizens support an expensive health care system for those near the end of life who have progressive, degenerative, and incurable conditions and will never return to economically productive lives? At the very least, solidarity and related strong notions of ethical recognition of membership in a human moral community and the mutuality of care, dignity, and respect give us important moral ideas with which to answer that question. I turn now to one way to think about what solidarity means and what practicing it involves. We can distinguish among different moments or stages in the practice of solidarity as it engages with moral recognition and mutuality. I suggest that the fundamental gesture and stance of solidarity are standing up beside another, thereby signaling publicly one’s recognition of that person’s (or group’s) moral standing or considerability. In addition, there are two distinct modes of mutual relationality that follow from this recognition, and that are particularly germane to end-of-life care. These may be called standing up for and standing up with.

Standing up for. Solidarity as standing up for is advocacy. It involves assisting, defending, and pleading the cause of the other. In general discussions of solidarity, the other is often taken to be a stranger, but that is not necessary; it makes perfect sense to talk about being in solidarity with one’s friends, although in this case it may well be that we build the concept of solidarity into the concept of friendship without noticing what we are doing or commenting on it. Moreover, the other need not be a human individual: one can stand up for other species, an ecosystem, or a cultural way of life. What is crucial is that there is some kind of power or knowledge differential between self and other in a relationship of solidarity. Otherwise the doing-for aspect of the relationship would be less salient. However, there is a serious shortcoming inherent in standing-up-for solidarity. It can assume and perpetuate a continued differentiation, even condescension, in the relationship between those individuals or groups standing up and those on whose behalf the standing is done. This kind of solidarity can advocate for improved treatment or benefits for an oppressed or vulnerable group, but does not necessarily challenge the underlying basis for their subordinate social status. Solidarity as standing and doing for others can become self-perpetuating. Inasmuch as it does not undermine structural inequalities, it serves to perpetuate the need for its own continuing existence.

Standing up with. Solidarity as standing up with takes another step in the direction of mutuality and recognition of shared moral standing. Moving from a mode of mutuality for to mutuality with requires deeper engagement with the experience and lifeworld of the other person. This closer understanding can tend to humanize and personalize the relationship further, and the potential benefit becomes more reciprocal. Those who stand with others may find their own initial pre judgments concerning
others transformed by the encounter. Solidarity can be stopped in its tracks and thwarted by fear of such self-transformation. Those who practice solidarity must open themselves to the opportunity—and risk—it poses. Relating to other people or groups in the specificity of their values and vocabularies of self-interpretation can simultaneously develop respect for the specific standpoints of others. It can also enable a greater capacity for intercultural and transpersonal interpretive understanding. Solidarity contains the possibility of being common readers of the diverse and distinct lives we each author.

In the end-of-life care context, an example of this transformation often arises in hospice programs that make room for dialogues of memory and reconciliation between a dying person and members of their family and friends. Family members often learn things during the hospice caregiving process that they did not previously know or understand about a loved one. Having the opportunity to discover this can be of lasting benefit after the person’s death.

If relational autonomy requires respect for the personhood of each and the moral membership of all, and if the practice of solidarity in the modes of standing up for and with involves a deep understanding and interpretation of a lived experience perceived through a distinct cultural and personal lens, then autonomy and solidarity find their point of convergence and reciprocal support. Solidarity involves finding a kind of covering connection that does not negate the autonomy of either oneself or the other with whom one stands in solidarity. Instead, it establishes the grounds of respect for and the social perpetuation of autonomy. Reconceiving autonomy in relational terms and bringing it together with solidarity allows us to say that we stand here both as human beings made capable through the functioning interdependence of a complex social community and as human persons with rights, needs, and vulnerabilities common to our kind, our humanity. One of these vulnerabilities is mortality; another is suffering. As Dame Cicely Saunders, describing St. Christopher’s Hospice in London, aptly put it, solidarity and autonomy are found in “the common giving of people who share the cost of being vulnerable”.

**Standing Alone before Death: Empowering the Autonomous Choices of the Dying**

Beginning in the early 1970s, innovations in ethics, law, and medicine began a movement toward empowering dying patients and their surrogates to have a greater voice in medical treatment decision-making, particularly in the face of progressive, chronic, and ultimately fatal disease. The landmark New Jersey Supreme Court decision (In re Quinlan 1976) that permitted the removal of a life-supporting mechanical ventilator from Karen Ann Quinlan—a young woman who had suffered catastrophic anoxic brain injury some years before and was being kept alive in a permanent unconscious state—gave this movement tremendous momentum. Over time, the
authoritarian, albeit benevolent, paternalistic decision-making style of many physicians was called into ethical question and came to be supplemented—if not entirely displaced—by the rights of patients to determine the nature of their medical treatment in accordance with their own preferences and values. In a growing number of court decisions, a new body of law established the legal right of patients to informed consent and to informed refusal of even life-sustaining medical treatment, using what was often called a “right to privacy” or a fundamental “liberty interest” protected by the constitution. State legislatures followed suit by enacting statutes reinforcing these rights and providing means, such as advance directives, for realizing them in practice. In ethics writings and in standards promulgated by various professional bodies in medicine and nursing, this kind of recognition and empowerment of the voice of the patient was expressed as the ethical principle of respect for autonomy or self-determination. What was meant, at bottom, by “autonomy” in bioethics at that time was an individualistic idea of liberty as noninterference and respect for self-sovereignty.

Being respected by others as autonomous was taken to be an intrinsic value in its own right, but the patient empowerment strategy was also oriented toward goals of beneficence and improvement in the quality of care and the quality of life experienced by the patient and others during the dying process. Paradoxically, changes taking place at about this same time in medical science and technology were making it possible to sustain vital function and to prolong biological life much more effectively than before. However, there was a widespread perception that due in part to this very technology, the dying individual was being objectified in the medical setting, and that physicians, now preoccupied with the technical challenge of intensive, invasive procedures, were losing sight of the whole person. Medical treatment was becoming divorced from medical care. Under these circumstances, relying on more or less unilateral physician expertise and paternalism was no longer ethically trustworthy or practically adequate to ensure a good dying marked by dignity, respect, relief of suffering, and human presence and witness.

Can life-extending medical technology be deliberately withheld or withdrawn when the patient refuses to consent to those treatments or when they offer no benefit but serve only to prolong the patient’s dying? There was a time when a physician was on solid ethical ground in doing “everything possible” to stave off death. The person confronted with a life-limiting illness was seen as the victim of an attack who enlists the knowledge, skills, and technology of physicians’ medicine as the defender and rescuer. This image of the physician’s role in the face of impending death can certainly be seen as the solidarity of standing up for—the solidarity of rescue, so to speak. Solidarity also provides one potential way to understand the role and responsibility of a health care surrogate or proxy in end-of-life decision-making.
Moreover, the situation of end-of-life care poses an ethical challenge for the concept and practice of solidarity in an especially poignant way. In many instances, medical decision-making near the end of life presents a tension between the duty to respect individual autonomy and the duty to protect and promote the best interests of the dying person. At stake is the timing of a person's death and the manner of his or her dying. If a patient's stated wishes or advance directive calls for treatment that is judged to be futile and unduly burdensome, a dilemma is created for surrogates, family, and physicians. The same problem arises if the patient's wishes indicate forgoing treatment that others believe would be beneficial. I think that the concept of solidarity per se does not resolve this ethical tension in end-of-life care or elsewhere. It does perhaps cast the tension in a new light and a new sensibility, however. The call of solidarity to stand up for a vulnerable other seems naturally to suggest primary attention to the best interests of the other, and hence there is a tendency for this mode of solidarity to embrace a paternalistic stance (if by paternalism we mean overriding subjective patient will and choice in favor of the patient's objective best interests). There is no logical or necessary reason why this should be the case, for there is nothing in the meaning of the concept of solidarity as I understand it that would preclude being in solidarity with the autonomy of the other. Nonetheless, it has been my observation in clinical settings that solidarity with the dying often means protecting them from something, and that something is most often neglect or nontreatment, rather than the use of interventions and treatments, which, in truth, are often actually abusive in their effects.

In any case, the distinction between solidarity as standing up for and solidarity as standing up with can be helpful here. The closer degree of imaginative connection with the other as a unique person that comes during standing up with acts as a counter to the advocacy mode of standing up for and protecting. It leads one toward viewing the situation through the eyes of the other—the internal properties of the situation, so to speak—rather than focusing on the objective external properties of the other's situation. Perhaps the challenge of the tension between the duty to respect individual autonomy and the duty to care for the individual's best interests is not so much to resolve it as to keep in view simultaneously the external consequences and the internal meaning in a person's life; in other words, to adopt both a third-person and a second-person standpoint in ethical judgment. That is the zone in which solidarity does its conceptual work. And all practices of solidarity structure experience and agency in particular ways.

Undergoing treatment for a serious illness is very much like entering an alien world, a kind of despotic, totalitarian world in which one must suspend one's normal rights, freedoms, expectations, and routines and then submit to physical and emotional handling that under nonmedical circumstances would surely be condemned as abuse.
Acute medical treatment is also akin to social dislocation—leaving one’s normal social space and entering an alien world—because it involves the abandonment of one’s personal identity and everyday roles and the assumption of a very different persona and role. This experience can precipitate a feeling of loss of self-identity. It can cause an alienation from one’s past—and perhaps even from what one fears will be the future—when it is precisely that past that one is desperate to regain and the avoidance of that future on which the entire treatment enterprise is predicated. It can involve a detachment from one’s family and closest loved ones, who, still healthy and whole, reside on the other side of the looking glass.

Paradoxically, it was medicine’s success rather than its failure that called such a passive patient encounter with medical expertise and intervention into serious question. Beginning in the 1960s, powerful new drugs and medical technologies were prolonging life-functions in critically and terminally ill patients—mechanical ventilators, cardiac interventions, kidney dialysis, and the fine-tuning of vital functions made possible by intensive care units. It soon came to pass that, for perhaps the majority of patients at the end of life, the exact timing and circumstances of one’s death were not beyond anyone’s control but were in fact the result of specific decisions and choices made by those caring for (physicians) and speaking for (usually family) the patient. Gradually, thoughtful leaders in medicine, and many family members who had observed the new end-of-life care firsthand, became aware that the experience of dying (for the individual, for the family, and also at times for health care providers) was often a horror. Few deaths were easy or “natural”; most occurred only after a serious of invasive and burdensome interventions, well-intentioned no doubt, but nonetheless very deleterious to the dying person’s quality of life and disruptive to connection and communication between the patient and family during the final weeks or days of life.

In sum, while life was being prolonged, many were living through an avoidably bad dying—a dying with inadequate relief of pain and suffering, inadequate compassion, inadequate human presence and witness; a dying marked by fear, anxiety, loneliness, and isolation; a dying that effaced dignity and denied individual self-control and choice. How did this happen? Powerful new technology was not being used wisely and judiciously but had succumbed to a kind of technological imperative. Aggressive interventions were becoming the default in hospitals and sometimes pushing the patient’s own wishes and values to the margins of decision-making. Reductionistic medical perspectives, which were the flip side of the new technology and the medical training that went along with it, were displacing more holistic perspectives, and this made it more difficult to see “the patient as person,” as one leading theologian-bioethicist aptly put it.

By about 1990, this situation was changing, although the forces of technology and patterns of inadequate palliative care and poor advance care planning persist
today. In the 1990s, the right to refuse unwanted and burdensome life-sustaining medical treatment was solidly established and the ethical obligation to respect patient autonomy was widely recognized in mainstream medicine. Nonetheless, the patient empowerment strategy of the past forty years has not been entirely successful. The shortcomings of the patient empowerment strategy take several forms, and they have been summarized and extensively documented in two successive studies by the Institute of Medicine.

Technologically driven orientations remain powerful in medicine and physicians do not always accept or cooperate with patient and family wishes; or else they influence and shape those wishes unduly.

Family dynamics and emotions complicate the identification of a patient’s autonomous preferences and values, and hence make it unclear what precisely others are obliged to respect when they respect autonomy.

And cultural or religious traditions and ways of understanding right conduct in the face of illness and dying have often not embraced the seemingly individualistic orientation of the concept of autonomy and the underlying concern for power and control implicit in the patient empowerment reform strategy in end-of-life care.

Standing Together before Death: Enabling the Relational Judgment in the Lifeworld of Dying

Prompted by problems such as these, another new reorientation—more focused on relationality, communication, and shared decision-making—is getting underway. There are several facets to this relational approach. It emphasizes shared, informed medical decision-making and consensus building among providers, surrogates, and family members, and the dying person, if capacity has not been lost. Attempts are also being made to improve continuity of communication and care planning among rotating providers and different health care settings. Staff training resources are being developed to improve communication styles and messaging so that they are culturally sensitive and competent.

Changing processes and practices such as these may be characterized in the phrase “Reconstrue rights, work through relationships.” Ethical and conceptual clarity will be needed to guide this shift if it is to realize its full potential, and also to protect against the undue erosion of important patient-centered rights and values. It is important to emphasize that while the rights of dying patients should be reconstrued, they should not be rejected. And it is especially important to insist on this distinction in the United States today, for this is a time of concerted, often less than careful or discerning financial costcontainment policy measures across the board that can have deleterious effects on end-of-life care.

The discourse of solidarity can be helpful in our social discussion of end-of-life
care at two levels. One is on the system and policy level, where solidarity, just
recognition, and just mutuality can appeal to rights and autonomy in a relational way
and provide justification and legitimacy to public expenditures, universal insurance, and
social investment in a humane and dignified dying process for all persons. The other is
on the more intimate level of care planning and decision-making at the bedside, where
a new understanding of the roles and obligations of surrogates and other family
members, as well as physicians and nurses, can be illuminated by seeing the situation
in terms of the trajectory of solidarity from standing for to standing with. Now
surrogates tend to be seen legalistically as advocates or formal representatives.

To develop this line of thinking further, we need to consider the substantive
content and purposes of dialogic communication and relational judgment in end-of-life
decision-making. I turn first to dialogic communication.

Relief of suffering, respecting individual rights, and preserving human dignity
and fulfillment at the close of a person's life are the core values that good
communication in the care process serves. Most dying patients—and the vast majority
of grieving family members too—do not actually desire medical treatments that are
known to be futile in advance or that most likely will leave the patient worse off, not
better off. Moreover, while prognosis concerning life expectancy and quality is still
difficult and uncertain, objective scientific information and tools are being developed
that enable physicians to make reasonable judgments about the effectiveness and
benefit of a given treatment to a particular patient. If patients, or more often, family
members for dying patients who have lost decision-making capacity request overly
aggressive and burdensome life-sustaining treatment interventions, it is not usually
because they have certain principled moral or religious beliefs of a vitalistic kind, but
more often because they do not accurately understand the nature and consequences of
what they are requesting.

Good communication is communication that will level with family members
about prognosis and treatment modality outcomes and keep them apprised in a timely
way of changes in the patient's condition so they have time to assimilate disturbing
new information emotionally before suddenly being called upon to make a wrenching
nontreatment decision as a surrogate. Good communication is also based on an
understanding of the social and emotional needs of family members and their cultural
worlds. In these worlds, kinship roles, spiritual orientations, and the responsibilities of
caregiving in the face of sickness and death all have a powerful meaning. We have
created a health care system in which medical choices are taken as the medium of
fulfilling these social and emotional meanings and goals. Fortunately, however, cultural
and personal meaning can be supported and respected without having to make
unreasonable decisions to request unduly burdensome life-sustaining medical
treatments.
Instead of solely focusing on how to empower the decisions of individual patients one at a time, we should ask how we can design a relational ecology of communication, recognition, respect, and good judgment in using technology and in meeting the needs of the dying patient as a whole person. I believe that the cognitive biases of all parties to end-of-life decision-making—providers, patients, and families alike—can be remediated by reflective judgment and participation in dialogic communities offering parity of voice and engagement to all members. Autonomy, human rights, and interdependent relationships of membership and mutuality are best understood as interwoven in the fabric of solidarity through which individuals acting together can each attain their own mode of flourishing and self-realization in ways that no one of them could attain by acting alone. That is the case in social life generally, and it is the case as well at the end of life.

How would viewing end-of-life decision-making through the prism of solidarity—that is, within an ecosystem of dialogic discourse and social practices that affirm moral recognition and mutuality—help preserve the ideals of personal autonomy and of respect for persons and their dignity? Social practices already permeate everyday life so it is not an architectonic question of building such forms of relationship and activity from the ground up. It is a more situated, practical question of mending or healing broken practices or sending out extensions of existing practices into new domains. In this way, I believe, we can arrive at a notion of autonomous medical decision-making as the exercise of relational judgment.

By relational judgment, I mean the capacity to discriminate among available courses of action on the basis of an interpretive understanding of shared values embedded in an ongoing institutional practice and in a broader form of communal life. Judgment is context-sensitive. It does not seek to make global and ultimate determinations but rather provisional, local ones—decisions for the here and now. In saying that it involves discrimination among available options, I mean to underscore the fact that the options are immediate, tangible, doable. Since the options are real, not hypothetical, and since the social world normally offers a rather narrow range of effective options at any time, the options are not usually grossly different from one another. They present instead subtle, fine-grained differences of meaning and implication. This is manifestly the case in end-of-life care. Catching the nuance, “discriminating” aptly, is highly sensitive to present attitudes, understandings, and probable reactions by others.

End-of-life care has seen professional paternalism, which is indeed imperious and controlling, albeit benevolent. It has seen individualistic autonomy, which has been very difficult for patients and families to sustain. Relational judgment is an alternative to both of these. It is an enmeshed practice of nurture, protection, provision, care, and support for another person. Relational judgment does not override autonomy, but
contextualizes it by giving the individual receiving care a new insight about his or her own agency, a new interpretation of the intentional and expressive possibilities that are actually present in a given situation, even the situation of curated caring during the time of dying.

Judgment is fundamentally dialogic. It requires a space for open dialogue and room for maneuver. Particular interpretations of the values embedded in a common form of life can always be challenged as misleading or lacking in depth or insight. One person's judgment and reasons in action are not totally incomprehensible or inaccessible to others, nor need the judgment of one person be a foreign or wholly extrinsic imposition on others. Judgment, in the sense I intend here, is not imposed but rather imbibed. It is absorbed and shared in a dialogue of question and answer among persons who are autonomous in the sense that they are indeed the authors of their own acts and lives, but they are symbiotic, not solipsistic authors. This agency of relational autonomy informed by judgment is a matter of mutual adjustments, mediations, trial and error innovation. It is an autonomy that savors the shared engagement of reflection and learning for its own sake as much as the attainment of its ends or the granting of its claims.

When autonomy is respected and practiced within a framework of reflective judgment in this sense, it will be open to reconciliation with caring counsel or intervention, just as caring practice will be open to respect for autonomy. The obligation to respect the self-determination of a relational agent will embrace the obligation to provide care and to promote the person's well-being.

**Conclusion**

There are some fundamental aspects of the human condition and human experience that naturally lead us up a path of increasing moral awareness. In the social world, solidarity traces one such path; we grow in moral awareness as we come to recognize the moral standing of others to whom that standing is being—or could easily be—denied, and then as we move from understanding ourselves as in relationship for others to being in relationship with them. That means that we not only are willing to support and defend others but also recognize their vulnerability as our own. We grow to understand that linkages of interdependence and symbiosis are not exceptional but pervasive and essential in human life. A serious reflection on death and dying—and even more so having gone through the lived experience of caring for a dying person and working to support continued human flourishing to the end—is an occasion, some would say the foundational occasion, for this sort of moral growth and learning.

Those who are dying are among the least valuable assets in a society limited to instrumental cooperation and individual self-interest. When the conversation centers,
as it so often does these days, on economic value and finding the most cost-effective ways to provide care to persons in the last year or few months of their lives, the equally important question of the moral value of such care to individuals, families, and society as a whole gets swept under the rug. Many would like to maintain that the ethical justification of high-quality end-of-life care is based on its goodness (the facilitation of human flourishing) and rightness (equitable respect for the dignity and worth of all persons). The economics of that care, real and valid though it is, is but a prudent means to those ends. Yet it is difficult—perhaps impossible politically if not logically—to sustain this argument without recourse to the concept of solidarity.