

8 Responsible Risking, Forethought, and the Case of Human Gene Editing

Madeleine Hayenhjelm

8.1 Introduction

Is there such a thing as “responsible” risking? Risky decisions are often colloquially criticized for being “irresponsible” and various morally attractive approaches to risks could plausibly be described as “responsible.” A recent example of this normative use of “responsible” can be found in the debate on the controversial issue of human germline gene editing: the possibility to make heritable changes to human DNA in embryos made possible by CRISPR.¹ For example, the high-profile names of the organizing committee of the gene editing summit made a statement in 2015 that it would be “irresponsible to proceed” with germline gene editing unless safety issues had been resolved and public consensus had been achieved (National Academies of Sciences, Engineering, and Medicine 2015). When Jiankui He in 2018 announced that he had edited the genome of two twin girls who had just been born, this was quickly condemned by a unified scientific community and criticized for being “irresponsible” (National Academies of Sciences, Engineering, and Medicine 2018).

What is interesting about the notion of “responsible risking” is that it points toward a potential middle category of moral advice: somewhere between advice based on moral certainty and advice based on measures of mere precaution. Or, so I shall argue. It does not require moral certainty about what is the morally best thing to do objectively all things considered. It merely requires that we do what is the wise thing given what is known and what can be done. If this is right, then it could provide action-guidance before we know for sure what is and is not morally right to do, and it could also provide action-guidance when many actions are permissible but not all are “responsible.” Additionally, it could provide a measure of caution that would be less permissive than standard utility maximization but more permissive than some versions of the precautionary principle. If there is something to this idea, and the idea of responsible risking is not just trivial or redundant, it could expand our toolbox for moral advice

in interesting ways. Given the uncertainties that come with moral uncertainty on the one hand and the uncertainties about risks on the other, this could turn out to be very useful.

Even if there is a general case for responsible actions, there may not be a case for responsible *risking*. To act in risky ways or impose risks on others is to act in ways that could go wrong, could cause harm, and could cause damages. This is what risking means. All of this seems contrary to “responsible” actions that imply some level of caution and forethought. In fact, there is a debate that argues that we have a right not to have risks imposed upon us (McCarthy 1997; see also, e.g., Hayenhjelm and Wolff 2012; Holm 2016; Steigleder 2018). Thus, there is at least room for an argument that if we violate, or at least infringe, the rights of others when imposing risks upon them, then it is also irresponsible to do so. On the other hand, if risk impositions sometimes are morally permissible, and it seems hard to avoid that conclusion if we want to avoid a problem of paralysis (Hayenhjelm and Wolff 2012), then it seems sensible that morality would ask us to impose such risks with great care, forethought, and in ways that we would generally describe as “responsible.” It seems wise to only impose risks with a certain degree of constraint, to not impose risks lightly, recklessly, or carelessly. Acting “responsibly” comes with precisely such connotations: to impose a risk “responsibly” implies that we proceed on good grounds, in careful ways, and are cautious not to bring about unnecessary harm. Thus, there seem to be enough intuitive grounds for a notion of “responsible risking” such that it is worth investigating the matter further.

I shall in the following address this question: Is there a case for responsible risking as a normative fruitful concept that could provide moral guidance when it comes to risk impositions? The main claim that this chapter will defend is this. Responsible risking already entails something of a forethought condition that would require a person to think ahead and try to anticipate future events. Furthermore, there is substantive moral content in the relevant notions of responsibility that could translate into moral requirements. If we pair these two ideas together, the forethought condition and the latent moral content in requirements from responsibility, we get three moral constraints that jointly give an idea about what would and would not be the responsible thing to do when imposing risks.

The chapter is structured as follows. The first section will set the stage and say something about key terms and main assumptions for the inquiry. The second section introduces the Forethought Condition. The third section turns to different notions of responsibility relevant to our inquiry and obligations from responsibility. The fourth section translates these obligations of responsibility into moral constraints on responsible

actions. The fifth section tests the three conditions on the example from human gene editing. Finally, a last section sums up the main points and conclusions.

8.2 Preliminaries

First, a note on terminology. By “risking,” I mean a deliberate action that may bring about unwanted harm as one of its (reasonably direct) outcomes (cf. Hansson 2004; Oberdiek 2017). Here, the focus lies on risk impositions – that is, risking that imposes a risk of harm on others. Thus, I impose a risk on you if and only if I act in a way that introduces a new source of harm (that may or may not result as a direct consequence of my action), or I act in a way that increases a risk of harm (that may or may not result as a direct result of my action), or in other ways that contribute to or increase risk of harm in some more indirect way.²

“Responsibly” refers to an evaluative and prescriptive notion beyond descriptive notions of responsibility. It is here understood as a thick moral concept. To act and risk responsibly is to live up to some substantive idea about being responsible and acting in a way that morally responds to descriptive responsibility in a morally good way.

By “responsible risking,” I mean risking that is performed in a manner that we would consider responsible and “responsible” is understood as a thick moral concept. In other words, we act responsibly if we act in a way that a responsible person would have acted or ought to have acted.

Next, a note on temporality. Risking is about what has not yet happened. To impose a risk is to act in a way that could cause harm in the future. This is what makes the action risky, but whether it in fact will cause harm is not known at the time of action. By contrast, liability, blame, reparations, answerability, and similar notions of responsibility refer to what has already happened as a result of someone’s actions – and often in light of what was known at the time of action. When we are inquiring into a notion of responsible risking, we are looking for a way to impose risks in a morally decent way that will align with the fact that we will be responsible for the outcomes but cannot be certain about what will be.

Both risk impositions and moral responsibility move across three nodes in time: (T1) the time of decision, (T2) the enduring phase of the action set in motion and the time when complications may arise, and (T3) the time after the effects. For some actions, T1 and T2 may be almost identical, as when knocking over a vase. For other actions, such as long-term policy decisions, T2 can extend over decades or even centuries. Part of the challenge of making a sound decision at T1 is that what seems wise to

do at T1 may no longer seem that way at T3. If we may make the wrong decision at T1, we may not be able to give any satisfactory explanation of our actions at T3.

The notion of “responsible risking” needs to operate across all three nodes of time. A person acts responsibly only if they at T1, as far as reasonably possible at the time, properly consider and prepare for what could arise at T2 such that they can reasonably handle relevant situations that may arise, and properly consider and prepare for what may be the result at T3 such that they can accept responsibility for those outcomes and, as far as reasonably possible, take on the obligations that follow from those outcomes.

8.3 The Forethought Condition

Lucas (1993) provides a perfectly good illustration of what this temporal component might look like in terms of responsibility as answerability:

If I accept that I may be legitimately questioned, I shall have that possibility in mind, and consider what answers I should give to questions that may be asked of me. I shall think about what I am doing, rather than act thoughtlessly or on impulse, and act for reasons that are faceable rather than ones I should be ashamed to avow.

(Lucas 1993, 11, §1.5)

There is a stroke of genius in the above quotation. Normally, we tend to think about backward-looking responsibility as something that begins after the consequences are known: Why did you do this? What can you do to fix this? Similarly, we tend to think about forward-looking responsibility, such as role responsibility, as beginning at the point when we deliberate upon our actions as choices about what to do from then on. After the consequences are known, we hold persons to account, perhaps demand reparations, explanations, apologies, etc. Before actions are embarked upon, we tend to think about what reasons, at the time of decision, would make sense to all affected as the situation is then understood. Here, these two ideas are combined: we ought now (before the action) to consider what reasons would and would not make sense in a future scenario that could result from our actions.

We could, as part of our deliberation, planning, and acting, anticipate how our current reasons may appear to those affected by our actions should things go wrong. The notion of risk and epistemic risk complicates this picture, given that our actions may come to affect different persons other than those that we had in mind at the time of deliberation and planning, and they may require a different kind of answer. Specifically,

that it never occurred to us how certain minorities would be affected by our actions would not be a satisfactory answer. In general, acting on insufficient knowledge, when relevant knowledge is hard and expensive to come by, may seem like a good idea before proceeding but serve as a very poor excuse should people come to serious harm or be discriminated against.

Hansson (2007), in a paper on risk impositions, argues that any decisions about risk ought to be guided by a kind of “hypothetical retrospection” that follows strikingly parallel ideas about forethought. When contemplating what to do, we ought to put ourselves in the hypothetical shoes of our future selves looking back at our action for each possible sets of consequences. If we pair this idea of going over all the relevant outcomes in our mind and viewing them with hypothetical retrospection, we must also keep in mind that different outcomes could affect different people, and they may require different kinds of answers. The relational aspect to answerability (see, e.g., Duff 2005; Gardner 2003) thus adds an important but challenging aspect to the forethought idea.

Both Lucas and Hansson, albeit in very different ways, point to a kind of responsible action (although Hansson does not use that word) where we *now*, before the action, need to try to imagine what kind of position we would be in, in the future, if we were morally responsible for that action and this had already occurred, and let that foresight guide our actions. In other words, to act responsibly is to anticipate what one may owe to others as a consequence of one’s actions before the consequences have occurred or the action has been taken. Even if we do not know what will happen but only that things could go wrong, we can always come more or less prepared for such outcomes. One does not travel to the Himalayas without preparations. However, we will not only need to come prepared for emergencies but also for the fact that we will be the ones in charge and the ones to face those that potentially come to harm. This would imply a different kind of preparation: in the form of reasons that could justify our actions and means to repair things should they go wrong.

Perhaps we could refer to this basic idea as *the forethought condition* of responsible risking. As a point of departure, it might look something like this.

The Forethought Condition: to act responsibly, one must deliberately plan and act in a way that is compatible with later being able to deliver on one’s obligations from responsibility over those actions and plans.

Stated thus, the Forethought Condition seems like a simple enough idea and relatively plausible as a core idea about responsibility. However, as

with any simple idea, we need to make it a bit more precise before we can assess its implications.

One source of ambiguity is that the forethought condition as stated is vague about the temporal aspect. The obligations from responsibility at T1, T2, and T3 are not the same. However, all decisions made at T1 and T2 could potentially be relevant at T3 in terms of explanations and answers. Thus, in one reading, the obligations of primary interest would be those at T3 given that the extent to which these can be satisfied will depend upon what was done and not done at T1 and T2. If we have failed to deliver on our obligations at T1 or T2, we will, most likely, be answerable for this at T3. To make the temporal aspects more explicit, we could rephrase the forethought condition in the following way and limit our investigation here to what is responsible to do at T1 in light of one's duties at T3.

The Forethought Condition, version 2: to act responsibly at T1, one must deliberately plan and act (at T1 and T2) in a way that is compatible with being able to deliver on one's obligations from responsibility over those actions and plans at T3.

Another source of ambiguity is the vague wording in terms of moral requirements: to plan and act in a way that "is compatible with" later being able to deliver on one's obligations. More worrying than vagueness, however, is that as a general condition for responsible action it may be much too weak, and as a condition for risk impositions it may be too demanding. How strong or weak it in fact is will of course depend on how "obligations from responsibility" is to be understood. This will be discussed in the next section. Here it suffices to know that there are at least three such obligations that arise from being the person in charge; those that arise from being answerable to those affected; and to those that arise from being responsible to restore harms, losses, damages, and injuries. On the one hand, we can do many reckless things while not exactly excluding the possibility of being able to deliver on our obligations. We do not exclude the possibility of calling the fire brigade even if we start the fire by intent or accident. However, we may not be able to justify our actions, and we may not be able to repair the damages. Certain actions are not compatible with being able to provide any reasonable justification later. Thus, in combination, the three conditions may get responsible actions roughly right. However, once we turn to the matter of risk impositions, even a rather minimal account that merely requires that we do not directly undermine or foreclose our future capacity to deliver on our obligations from responsibility and are able to justify and rectify our actions may quickly become very demanding.

The challenge arises with the fact that when it comes to risk impositions, we cannot count on any one outcome as being certain but need to take all reasonably relevant possibilities into account. Furthermore, we may end up responsible for a situation we did not predict. We cannot responsibly act merely on the basis of the outcomes that we expect to occur. Many of the most damning risks introduced were not thought to be dangerous at the time of their introduction. This goes for carbon emissions, DDT, tobacco, amphetamines, plastic waste, etc. It does not seem right to thereby simply assume that because they were not foreseen, they could not have been irresponsible to introduce. Still, to claim that a risk imposition is responsible only if it is compatible with being able to deliver on obligations from responsibility across *all* possible scenarios, whether foreseen or not, seems too demanding. This would imply that we could never risk bringing about something that we later realized we could never reverse or repair or explain to those affected. Given that some risk policies could be in place over a very long time and that people could be affected in very indirect and unpredictable ways this could make any radical change, even for the better, irresponsible. Thus, we must limit the number of relevant scenarios to those that are “reasonably morally and probabilistically relevant.” This, again, is a vague idea and merely points out the direction (see Oberdiek 2017 for an interesting attempt to narrow down morally relevant probability).

The Forethought Condition, version 3: to act responsibly at T1, one must deliberately plan and act (at T1 and T2) in a way that is compatible with being able to deliver on one’s obligations from responsibility across all reasonably morally and probabilistically relevant outcomes at T3.

If we rephrase it in this way, the core idea becomes clear. To act responsibly, we will need to consider ways in which our actions may go wrong and have preparations for this, both in how to deal with it as it happens and in terms of fixing things and being able to explain our actions and decisions afterward. Ideally, we would be able to predict all ways in which things could go wrong, prepare for those outcomes, ensure that they would not arise and, if they were to arise, ensure that they would be properly taken care of. This is, however, unrealistic. When it comes to risk impositions, the actual risks are sometimes only learned about later. This means that for all new sources of risks, we may not even know what the relevant risks are. Thus, even the best decisions at T1 may leave one unprepared for T2 and without any good explanations at T3.

However, we need not know the precise risks in order to plan for eventualities. We may not know before the clinical trial whether a new drug is in fact efficient and safe, but we can know if the research subject

has given their informed consent, whether there are medical resources available to treat side-effects, and whether there is insurance to cover unwanted outcomes, etc. When it comes to being in charge of dangerous policies, experimental practices, and other kinds of gambles that could affect others badly, such plans may be essential to what is required from responsibility. This need not apply in the same way to what is required from responsible actions in the context of a normal person going about their normal everyday affairs that still pose some level of risk. Part of this is meant to be covered by the limitation to “probabilistically and morally relevant outcomes.” What in fact is probabilistically and morally relevant may be determined by moral norms of expectation – precisely of the kind that would determine when an explanation is sufficient to justify a course of action. We are not required to have an ambulance on stand-by when airing and dusting books on a balcony or to have funds set aside for the eventuality of food poisoning before inviting friends over for a seafood dinner. Part of this may have to do with the fact that this is not part of what we expect from each other and thus not something we would need to explain. Most cases of everyday risks can be remedied by apologies, simple measures of repairs (in a non-financial sense), and lessons learned for next time around. They can thus be imposed in ways compatible with later obligations from responsibility.

8.4 Obligations from Responsibility

What are the obligations presumed to follow from responsibility? To answer this question, we will first take a quick look at what responsibility in a descriptive sense roughly entails and, only then, return to what this would imply for being responsible in a more evaluative sense. The assumption is that to act “responsibly” is to be responsible, in a descriptive sense, in a good way.

Responsibility in a descriptive sense can refer to a number of different things. For example, the following four questions are all answered by a different concept of responsibility. (1) Who is in charge of A (where A is some action or domain)? (2) Who is to blame for O (where O is some outcome of an action)? (3) Who will fix O or compensate for O (where O is some outcome of an action)? (4) Why did you do A (where A is some action)?

There are various names attached to each of these categories. I will opt for role responsibility to refer to the kind of responsibility that answers the first question, blameworthiness for the second, responsibility to repair for the third, and answerability for the last one. For our purposes, we will

focus on 3 and 4 and to some extent on 1 but not further discuss blame-worthiness here.

Role responsibility. Who is in charge? The first concept would refer to a person being responsible by taking up or holding a position of authority over some domain or by being the moral agent who decided to act in a particular way. It connects an action to a moral agent. It could be forward-looking as in “Who will be in charge of A?” Backward-looking, as in “Who was in charge of A?” Or refer to an on-going position of responsibility over something or someone, as in “Who is in charge of this?”

Blameworthiness. Who is to blame? The second concept refers to blame-worthiness, the person to praise or blame for some action. This could also extend to accountability or liability. This is a distinctly backward-looking notion of responsibility. This is also the classical notion of “moral responsibility” that dominates much of the literature on responsibility.

Responsibility to repair. The question “Who will fix this?” points to reparative responsibility, which is in a sense forward-looking, but after the consequences have occurred. It points to a role of being in charge, but for reparations rather than as author of the original action.

Answerability. The question “Why did you do it?” points to someone as being answerable for their action to others affected by it. All the first three questions can be answered by pointing to a particular person: They are responsible for A, they are to be blamed for O, and they are the one to fix O. However, the fourth question is second-personal, “Why did you do A?” is aimed at the moral agent pointed out by the other questions (cf. Darwall 2006).

The concept of responsibility as answerability has been developed in recent decades by a number of writers such as Duff (e.g., 2001, 2013), Gardner (e.g., 2003, 2008), Smith (2015), Shoemaker (e.g., 2011), and others. The basic idea is this: to be responsible is to be answerable, that is to be able to or even be obliged to provide an answer for one’s action (Lucas 1993, 5). Gardner departs from the same basic idea and distinguishes between two kinds of answers: justifications and excuses – both explained in terms of reasons.

Responsibility is what it sounds like: it is a kind of ability to respond. More precisely, it is the ability to explain oneself, to give an intelligible account of oneself, to answer for oneself as a rational being. [...] As a rational agent, one only has two ways of explaining oneself. The first is to offer a justification; the second is to offer an excuse (Gardner 2008, 123).

All four of these concepts of responsibility point to various aspects of descriptive responsibility: of what it means to be responsible for something. A person can be responsible in the sense of being the legitimate target for blame and praise. It can also refer to them being the person in charge, them being the person with obligations to repair or compensate for outcomes, or them being the person obliged to justify their actions to those affected by them. The latter three meanings of responsibility are of relevance in this context.

Responsibility in an evaluative sense would then refer to taking on, accepting, or living up to such descriptive responsibilities in a good way. Normatively, acting responsibly in this evaluative sense is also something we ought to do when faced with such descriptive responsibility. To act responsibly is to accept responsibility in a way that someone who is good at being responsible would act. I shall suggest that there is more moral content to this idea than one might first suspect.

Thus, to act responsibly is to act in ways, early on, that allows one to successfully oversee a domain over an extended period. To act responsibly is to act in ways, early on, that allows one to, later, successfully repair what one has broken or harmed. To act responsibly is to act in ways, early on, that allows one to, later, have good answers for why one did what one did. To act responsibly is to take charge of what needs to be done in light of what was done and what such actions resulted in.

This “early on” clause is essential. Just as “precaution” has an element of prevention of later harm, “responsibility” has a similar preventative element that involves planning and preparation for later events and outcomes. We want to hold people to account, to hear their explanations when things go wrong. To take responsibility is to stick around when things go wrong, to admit mistakes, and to seek to explain and repair them when they occur. To act responsibly is also to take measures and make plans such that negative outcomes are to a reasonable degree foreseen and avoided.

8.5 Responsibility Conditions as Moral Constraints

Let us return to the Forethought Condition. The general idea was this: in order to act responsibly one must seek to avoid doing, at T1 and T2, what at that point in time, we have reason to believe will foreclose our ability to be in a position to deliver our obligations from responsibility at T2 and T3. The gist of it is that we cannot claim to act responsibly and at the same time undermine our ability to do what is required from us as responsible agents.

At least three of the different parameters of responsibility above give us different kinds of failures when not satisfied: role responsibility,

responsibility to repair, and answerability. We could fail in our capacity to be charge of a situation that we have a role responsibility to be in charge of. We could fail in our capacity to “fix” what it is our role to fix. We could fail in our capacity to explain and provide reasons to those affected for actions that we have performed. What is of interest here is that these failings could provide us moral limits to responsible risking. To put it in the words of the Forethought Condition: some decisions and actions (at T1 or T2) may be such that they are incompatible with the capacity to successfully deliver on one’s role responsibilities, to repair things that have gone wrong (or resulted in harm or loss), or on one’s responsibility to provide satisfactory answers. If this is correct, then this gives us at least three moral limits to responsible acting from the abovementioned three kinds of responsibilities: role responsibility, responsibility to repair, and answerability.

These limits could be expressed in the following three conditions:

The control condition. We ought to act in ways that allow us to deliver on role responsibility and have control over our domain of responsibility and over outcomes.

The responsibility to fix condition. We ought to act in ways that allow us to deliver on our obligations to repair things that go wrong.

The responsibility to explain condition. We ought to act in ways that allow us to provide reasonable and acceptable answers to those affected.

The first condition would require us to act in ways that are compatible with remaining in control over what is legitimately in our domain of responsibility and enable us to deliver on decisions and actions justly required from someone in that position of responsibility. This is compatible with delegating jobs and passing responsibilities onto others. What it would rule out are various ways of not being in control while in the role of such responsibility, such as being drunk when editing genes, or withdrawing from such a role without any plan or measure for such responsibility to be delegated or taken over by someone else. Additionally, it would also rule out initiating processes that could quickly expand and become uncontrollable.³

The second condition would require us to not act in ways that would make us unable to repair, replace, or compensate for losses and damages that we bring about. Furthermore, it would positively require us to act in ways that could contribute to our ability to repair and compensate for possible harms, losses, and injuries that we may cause. In order to do this, we must anticipate in what ways and to what extent things could be harmed or lost.

The third condition would require us to not act in ways that would render us unable to sufficiently explain and justify our actions to those

directly affected by them. Furthermore, it would require us to act on reasons that would make sense and “speak” to those affected. In order to deliver on this, we must anticipate whom our actions may affect and what their crucial interests, rights, and values are.

These conditions tell us, if correct, how to impose risks responsibly: to only impose risks in ways that we can maintain control over, to only impose risks that we would be able to fix or compensate for, and to only impose risks that we could justify and explain to those affected. They also give us a hint about risks that may be, categorically, off limits. Some risks are such that they could never be, or hardly ever be, reined in if control was lost. Some outcomes are such that they could never be repaired if the worst came to be. Some risk impositions are such that they could never be justified. In all such cases, the responsible thing to do may very well be to refrain from imposing such risks.

8.6 Responsibility Conditions and Human Gene Editing

Let us try out our three conditions on the controversial case of human germline gene editing. What would responsible risking look like when it comes to gene editing and human germline gene editing? If we apply the three conditions above to gene editing and risk impositions, then we would get something like the following:

The control condition. Responsible risking requires us to only impose gene editing risks in ways that allow us to remain in control over the risks within our domain of responsibility.

The responsibility to fix condition. Responsible risking requires us to only impose gene editing risks in ways that allow us to deliver on our obligations to repair things that go wrong.

The responsibility to explain condition. Responsible risking requires us to only impose gene editing risks on grounds that we can justify to others especially to those who have a right to an answer from us.

The control condition implies that we could not responsibly impose risks that exceed what we could remain in control over within our domain of responsibility. What is implied by the control condition is something like the following: we can only responsibly put into motion courses of action that will remain largely controllable, such that it will be possible to make new decisions, change direction, etc., should new challenges, new facts, and the like arise. In the context of germline gene editing, should it become a legally permitted practice regulations, permits, licenses, professional codes of conduct, medical ethical approvals, etc., would in all likelihood help to ensure controllability. The main reason why Jiankui He

was widely condemned in 2018 was the fact that he went against scientific consensus, professional codes, and regulations (Krimsky 2019).⁴

There are, however, three issues that could make it difficult to edit the human germline while not losing control over the risks and future developments as a result of this.

First, the fact that germline edits affect future individuals and that these edits are heritable.⁵ The effects thus lie in the future and could extend far into the future. This means that risks could appear after the original decision-makers are gone. This need not be a worry; there are many ways to extend control responsibility across generations via regulation, institutions, and reliable processes for delegation and appointments of roles of responsibility. However, whether control and the ability to deliver upon obligations extend across generations will depend on the proportionality between capacity and size of the tasks. This could change dramatically across generations and, if much larger for later generations compared to the earlier ones, it may exceed what could be considered fair to pass on to later generations. Some decisions are easy to make (before more is known) but very hard to manage at a later stage. In the case of germline gene editing, various unknowns could make role responsibility much harder for later generations than earlier ones. We could imagine an edit that seems very promising but leads to cancer in a significant number of individuals. We could also imagine a case where radical changes to human DNA would make us much more vulnerable in a future where there are rapid and dramatic changes in the natural environment.

Second, epistemically, we do not have sufficient knowledge about effects. There are two challenges here. The first challenge stems from the fact that genes can have multiple functions. One and the same gene could thus be causal to one type of cancer and at the same time prevent another type. Furthermore, many of the diseases, conditions, and vulnerabilities that we may want to edit depend on more than one gene. This means that beyond remedying cases of severe diseases that depend on a single gene, many things could go wrong and have unwanted side-effects. The second challenge stems from the fact that some side-effects may only appear later in life or in the second generation born with edits. This means that it is difficult to gain the full epistemic picture before we, so-to-speak, try it out (Guttinger 2018). Even though it is perfectly possible, in some cases, to be in control and act responsibly when exploring the unknown, there is a limit to how far into the unknown one can venture responsibly. Part of being in control is knowing what one is doing and why. This requires some basic knowledge about the relevant outcomes. Without such knowledge, it is hard to see how we can make responsible decisions or have good plans in place.

Third, the most serious objections to germline gene editing are not about its impact on individuals but on society and what it means to be a human and regard others as fellow human beings. These kinds of objections are thus concerned about the course we would embark upon. This is an often-repeated worry that, once we embark upon germline gene editing for severe genetic disease, we will push the norm forward and lead to a slippery slope where the bottom of the slope would represent something like a *Gattica* or *Brave New World*-type dystopia (Baylis 2019; Evans 2020). The fear is that slippery slopes could lead to genetically divided societies that not only sort people according to distinct genetically determined classes but add a new genetically enhanced elite with abilities that go beyond what the best among us currently can be or achieve. This would not only drastically deepen current inequalities but also make them more permanent by having them written into our DNA. Other concerns about dramatic societal impacts are about fundamental human values lost, such as human rights premised on being “born free and equal” losing their foundation (see, e.g., Fukuyama 2003). Such developments could have large-scale impacts and develop in ways that could end up being uncontrollable. It is hard to see how we could maintain control responsibility if society or human nature is too radically or too rapidly changed.⁶

Heritability, considerable epistemic gaps, slippery slopes, and large-scale social impacts all raise challenges for the control condition – at least under current levels of knowledge. The key point is that we will be responsible for how things develop and not just for how we imagined them to develop, so we ought to be able to stay on top of that and make responsible decisions if we are to impose risks responsibly in this sense.

The responsibility to fix condition would require that we act in ways such that potential harms, damages, losses, and injuries are largely reversible, repairable, replaceable, or compensable. There are different ways of “fixing” unwanted outcomes. Should something result in an unwanted outcome, we could, potentially, reverse it, such that we are back where we were before we imposed the risk, or the outcome came about. Should we not be able to reverse it, we could repair whatever has been “broken” or, failing that, replace it with an equivalent. Sometimes there is nothing that can replace or repair something. In such cases, other things may balance the loss by offering something else that is even better. To see what this could imply for gene editing, we must first assess what could go wrong, and the nature of potential losses and harms.

First, we have the technical risks: off-target risks, unwanted but on-target risks, and mosaicism. The outcome of an edit, if proven to have unwanted side-effects, could in some cases perhaps be remedied with somatic gene therapy; but, when it comes to radical alterations, this may not be possible and will need to be “remedied” via “re-edits” of the next

generation of embryos. Still, what cannot be “fixed” by re-edits could possibly be compensated for if not too grave. Not editing genes could also cause risks of harm. There is thus a reverse case, where responsibility may require us to edit an embryo to relieve it from causes for future suffering – especially in a scenario when this is an accepted practice. Here again, it may be possible to remedy this by somatic gene therapy in some cases and compensate for failing to do so in others. (These cases are further complicated by the fact that the decision to have a child at all may be conditional on the possibility of editing their DNA when there is knowledge about risk for severe genetic disease.) In general, however, the harm to an individual that may result from germline gene editing seems hard to “fix”.

Second, we have societal risks, such as risks for discrimination, changed social climates and norms, the undermining of cultures, civilization and the like. These could prove even harder to “fix” given that they would require big shifts in society, and, sometimes, they may prove impossible to go back on if the shifts are too far-gone and too much has already been invested in them (Mariscal and Petropanagos 2016). Some changes to history cannot be reversed. We cannot, for instance, undo the industrial revolution. Thus, if we set in motion changes as radical as that, it may not be something we could “fix” if things went horribly wrong. However, developments could be controlled so as not to lead to such radical developments.

Third, we have the existential risk, such as the potential loss of humanity as a kind (cf. Annas, Andrews, and Isasi 2002). This may seem a highly improbable outcome. Nevertheless, should we somehow bring about the end of humanity as we know it, this would be very hard to “fix”. In fact, it seems likely that this would constitute what I have referred to elsewhere as a “genuine loss” of a valuable kind – i.e., a loss that could not be repaired, replaced, or compensated for (Hayenhjelm 2018; Hayenhjelm & Nordlund forthcoming). What is in dispute is whether the loss of humanity is to count as a loss or as a gain. The transhumanists tend to think that the extinction of humanity could be thought of as a gain if replaced by a better, new species: the posthuman (Bostrom 2005). However, it is doubtful that this could be *our* gain rather than a permanent loss for us (Agar 2010; Levin 2021; Porter 2017).

In short, there are many potential outcomes that we would not be able to fix. These include harm to individuals, radical changes in norms and attitudes, potential social costs including discrimination and new genetic castes, and so on. However, most of these depend upon germline editing used for enhancement purposes. If limited to prevent medical conditions alone, or only to prevent or enhance resistance to disease, many potentially irreparable risks are avoided. However, even medically motivated edits come with risks that are not reversible for the person, and making these

sufficiently safe may require trials across multiple generations and imposing risks that may not be reparable or possibly not compensable if severe enough.

Many outcomes can be repaired, losses replaced, harms healed, courses retracted, or at least compensated for. Some outcomes cannot be reversed, repaired, or compensated for – the losses are too great to make it possible to ever be outweighed (Hayenhjelm 2018). In such cases, it seems that precaution would be the right kind of approach. Some actions could render us without any means to ever “fix” the outcomes. If the consequences are severe enough, it seems that in order to be responsible, in the moral sense, to merely offer an apology or accept moral responsibility as in blameworthiness, will not suffice to make the wrong right or repair what was done (Hayenhjelm 2019). Thus, it seems that precaution and responsibility to “fix” point in the same direction: whenever there is a risk of harm or loss of such magnitude that it could not possibly be fixed, then, unless solid ground warrants exception, we cannot possibly impose such risks and be responsible while doing so. The responsible thing is to refrain from imposing such risks.

The responsibility to explain condition would require that we only impose such risks that we can reasonably explain and justify to those affected by them. We can only responsibly impose risks that we could justify and explain to others and, failing that, be able to offer some kind of excuse for.

What would and what would not count as a valid excuse or explanation in terms of germline gene editing? To a large degree, this depends on the degree of risks as well as intentions and reasons. There are, of course, obvious cases that could never be excused: such as willful and deliberate edits done with the aim of harming another person or for experimentation that would not be in the person’s own interest. But even well-meaning edits that turned out to be unnecessary, riskier than thought, only relatively valuable, etc. could be questioned by the person(s) so edited.

This could rule out things like “donor siblings,” or any kind of germline edits that were not for the person’s own good. It would also rule out unnecessary risk-taking and acting prematurely before the risks are known and prepared against. We could not responsibly impose risks on groups of individuals that we were unaware might be negatively affected. For example, that we never considered that germline gene editing might negatively affect those with disabilities or functional variations is not a good answer to them (Sufian and Garland-Thomson 2021). More than that, as the Lucas quotation suggests, we also hold those in charge responsible for the way things are done, planned, and prepared for, as well as the number of backup plans, emergency measures, and kind of skills, training, resources, etc. that go into a responsible, but risky plan. We can accept risky and novel projects, but not sloppy and ill-prepared risky projects. “I never

thought of that” is just as poor an explanation as “I never thought this would affect you,” at least when “you” refers to a relevant reference group at risk. Even though not all consequences may be predictable, responsible risking would require allowances for more risks than those known about and preparations for what is not known. Thus, we may not know about all kinds of dangers in a jungle before entering it, but we could prepare as well as possible for all kinds of dangers that we can imagine. “Why did you not pack a knife?” is a perfectly good reproach even if one could not list all the possible dangers that would require a knife.⁷ Responsibility does not require perfect prediction, but it does require some level of reasonable preparations that go beyond what is known based on what could happen in light of similar cases and relevant knowledge.

In many cases, the demands from control, reparability, and answerability overlap. What can be fully fixed will often be under some degree of control, and what can be fully fixed can often be excused. What could not be fixed and could not be controlled is also hard to justify. Should it turn out to be the case that experimentations with human nature came to irreparably wreck our species, undermine our core values (such as seeing the “humanity” in each other, human rights, etc.), or undermine our civilization, it is hard to see how we could fix or excuse such an outcome. This seems to hold for most “genuine losses” – in most cases, there is neither a satisfactory excuse nor a fix that could make things right. We simply took risks that were too large or acted when it was epistemically premature given the risks.⁸ This is most likely where the hard limit to what could responsibly be risked lies. It should be mentioned that the three conditions could also pull in different directions. For example, we could imagine a case where some risky activity is so important that we would be held to account for not pursuing it even if we could not guarantee that we would remain in control over the events that followed.

8.7 Conclusions

Is there such a thing as “responsible risking”? This chapter has explored the notion of “responsible risking” as a thick moral concept. I have argued that the notion can be given moral content that can be action-guiding and add an important tool to our moral toolbox in the context of risk impositions. To impose risks responsibly, on the view defended, is to take on responsibility in a good way. A core part of responsible action, I have argued, is some version of a Forethought Condition. Such a condition requires us to not make decisions or plans such that we cannot deliver on our *responsibility obligations*. The morally limiting features come from what must be the case in order to be able to deliver upon one’s obligations from responsibility. I have looked at three such notions: *role responsibility*,

responsibility to provide reparations, and *answerability*. All three of these hold implicit limits that can be translated into normative boundaries. I have called these *the control condition*, *the responsibility to fix condition*, and *the responsibility to explain condition*.

This general idea of responsible risking was tried out on the controversial case of human germline gene editing. From the control condition, we can conclude that responsible germline gene editing would require us to only impose risks in ways that allow us to remain in control over the risks within our domain of responsibility. From the responsibility to fix condition, we can conclude that responsible germline gene editing requires us to only impose risks in ways that allow us to deliver on our obligations to repair things that go wrong. From the responsibility to explain condition, we can conclude that responsible germline gene editing would require that we only impose risks in ways that we can justify to others and especially to those who have a right to an answer from us.

Are these ideas about responsible risking substantive enough to be action guiding? If so, is there anything fruitful here that is not merely trivial, redundant, or covered by the standard moral answers to risk impositions? The notion of “responsible risking” defended here points toward three distinct parameters to responsibility and thus three kinds of reasons that could support decision-making about risk impositions. Responsibility is not meant to replace other moral notions but supplement them, especially when we do not have full moral answers. We can act responsibly also when we do what later turns out to be the morally wrong thing. In fact, it is the possibility that we may do what we later could have reason to regret that makes responsibility an important notion. We can act in ways that allow us to have control over risky activities, we can be prepared to repair what could later occur, and we can act in ways that we are willing and able to explain to those to whom we may come to owe an answer. This, I have argued, would be a responsible case of risking under uncertainty or incomplete moral knowledge.

Acknowledgments

The author would like to thank the Marcus and Amalia Wallenberg Foundation. This research is one of the outcomes of the Umeå University-research project “The Rise and Risks of CRISPR/Cas9: Cultural and Ethical Perspectives on the New Gene Technology” lead by Prof. Christer Nordlund, funded by the Marcus and Amalia Wallenberg Foundation. The author would also like to thank the two dedicated editors for this volume, Adriana Placani and Stearns Broadhead, for their patience, careful close readings, insightful comments, and for the initiative of putting this interesting volume together.

Notes

- 1 In 2015, when it was first demonstrated in a laboratory to be possibly to apply the technology to human (non-viable) embryos, this resulted in a number of leading scientists and bioethicists and others to call for moratoria or more generally a “prudent path forward” given the high risk for off-target effects (that edits result in unintended mutations elsewhere in the DNA), unwanted on-target effects (unintended mutations at the target site) and mosaicism (incomplete edits such that some cells are and some are not edited in the intended way), and the ethical issues it raises. See, e.g., Lanphier et al. (2015) and Baltimore et al. (2015).
- 2 What is of relevance is both direct and some indirect consequences that are proximate enough. Here I have merely used “direct” to exclude the more far-fetched consequences. This is somewhat crude but sufficient for the purposes of this chapter.
- 3 The underlying intuition here runs counter to that of the Doctrine of Double Effect; the key point is not what you intend but what kind of outcomes you bring about that you could oversee and rein in if need be.
- 4 Outside the specific topic of germline gene editing, concerns related to what we have referred to as the control condition have been raised against bio-hackers experimenting with gene editing (largely upon themselves) and gene editing paired with gene drivers essentially making malaria-carrying mosquitoes infertile with unpredictable effects on the ecosystem.
- 5 The heritability aspect may also make attributions of blame difficult for those born with unwanted edits.
- 6 Not very surprisingly the debate on CRISPR has focused on drawing moral lines to keep the development safe enough to avoid worst outcomes but not so restrictive as to not allow medical progress. See Evans (2020) for overview on the debate.
- 7 The lack of explanation or excuse could cut two ways: we could end up in a position where we could not “give” any reasonable explanation to others, and we could (also) end up in a position where we had no excuse that we could accept ourselves given the outcomes. At the far end, we could end up having performed an act that was largely “unforgivable” – by our own standards, or by those affected, or by the larger moral community. It is likely that some risks that we could never repair we could explain (the reasons seemed good at the time), and that some risks that we could not justify we can still reverse or fully repair.
- 8 For more discussion on “genuine losses” and harder cases of risks, see Hayenhjelm (2018) and Hayenhjelm and Nordlund (forthcoming).

References

- Agar, Nicholas. 2010. *Humanity's End: Why We Should Reject Radical Enhancement*. Cambridge: MIT Press.
- Annas, George J., Lori B. Andrews, and Rosario M. Isasi. 2002. “Protecting the Endangered Human: Toward an International Treaty Prohibiting Cloning and Inheritable Alterations.” *American Journal of Law & Medicine* 28 (2–3): 151–178.
- Baltimore, David, Paul Berg, Michael Botchan, Dana Carroll, R. Alta Charo, George Church, Jacob E. Corn, et al. 2015. “A Prudent Path Forward for Genomic Engineering and Germline Gene Modification.” *Science* 348 (6230): 36–38.

- Baylis, Françoise. 2019. *Altered Inheritance: CRISPR and the Ethics of Human Genome Editing*. Cambridge: Harvard University Press.
- Bostrom, Nick. 2005. "Transhumanist Values." *Journal of Philosophical Research* 30 (Supplement): 3–14.
- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.
- Duff, R. Antony. 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press.
- Duff, R. Antony. 2005. "Who Is Responsible, for What, to Whom?" *Ohio State Journal of Criminal Law* 2: 441–61.
- Duff, R. Antony. 2013. "Relational Reasons and the Criminal Law." *Oxford Studies in Legal Philosophy* 2: 175.
- Evans, John. H. 2020. *The Human Gene Editing Debate*. Oxford: Oxford University Press.
- Fukuyama, Francis. 2003. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. London: Profile Books.
- Gardner, John. 2003. "The Mark of Responsibility." *Oxford Journal of Legal Studies* 23 (2): 157–71.
- Gardner, John. 2008. "Hart and Feinberg on Responsibility." In *The Legacy of H.L.A. Hart: Legal, Political and Moral Philosophy*, edited by Matthew Kramer, Claire Grant, Ben Colburn, and Antony Hatzistavrou, 121–40. Oxford: Oxford University Press.
- Guttinger, Stephan. 2018. "Trust in Science: CRISPR-Cas9 and the Ban on Human Germline Editing." *Scientific Engineering Ethics* 24 (4): 1077–96.
- Hansson, Sven Ove. 2007. "Hypothetical Retrospection." *Ethical Theory and Moral Practice* 10 (2): 145–57.
- Hansson, Sven Ove. 2004. "Philosophical Perspectives on Risk." *Techné: Research in Philosophy and Technology* 8 (1): 10–35.
- Hayenhjelm, Madeleine. 2018. "Risk Impositions, Genuine Losses, and Reparability as a Moral Constraint." *Ethical Perspectives* 25 (3): 419–46.
- Hayenhjelm, Madeleine. 2019. "Compensation as Moral Repair and as Moral Justification for Risks." *Ethics, Politics & Society* 2: 33–63.
- Hayenhjelm, Madeleine and Christer Nordlund. (forthcoming). *The Ethics and Risks of Human Germline Editing: A Philosophical Guide to the Arguments*. Springer V.S.
- Hayenhjelm, Madeleine, and Jonathan Wolff. 2012. "The Moral Problem of Risk Impositions: A Survey of the Literature." *European Journal of Philosophy* 20 (S1): E26–E51.
- Holm, Sune. 2016. "A Right Against Risk-Imposition and the Problem of Paralysis." *Ethical Theory and Moral Practice* 19 (4): 917–30.
- Krimsky, Sheldon. 2019. "Ten Ways in Which He Jiankui Violated Ethics." *Nature Biotechnology* 37: 19–20.
- Lanphier, Edward, Fyodor Urnov, Sarah E. Haecker, Michael Werner and Joanna Smolenski. 2015. "Don't Edit the Human Germline." *Nature* 519 (7544): 410–11.
- Levin, Susan B. 2021. *Posthuman Bliss*. Oxford: Oxford University Press.

- Lucas, John. 1993. *Responsibility*. Oxford: Oxford University Press.
- Mariscal, Carlos, and Angel Petropanagos. 2016. "CRISPR as a Driving Force: The Model T of Biotechnology." *Monash Bioethical Review* 34 (2): 101–16.
- McCarthy, David. 1997. "Rights, Explanation, and Risks." *Ethics* 107 (2): 205–25.
- National Academies of Sciences, Engineering, and Medicine. 2015. "On Human Gene Editing: International Summit Statement" (13 December 2015). <https://www.nationalacademies.org/news/2015/12/on-human-gene-editing-international-summit-statement>
- National Academies of Sciences, Engineering, and Medicine. 2018. "Statement by the Organizing Committee of the Second International Summit on Human Genome Editing" (25 November 2018). <https://www.nationalacademies.org/news/2018/11/statement-by-the-organizing-committee-of-the-second-international-summit-on-human-genome-editing>
- Oberdiek, John. 2017. *Imposing Risk: A Normative Framework*. Oxford: Oxford University Press.
- Porter, Allen. 2017. "Bioethics and Transhumanism." *The Journal of Medicine and Philosophy* 42 (3): 237–60.
- Shoemaker, David. 2011. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121 (3): 602–32.
- Smith, Angela M. 2015. "Responsibility as Answerability." *Inquiry* 58 (2): 99–126.
- Steigleder, Klaus. 2018. "On the Criteria of the Rightful Imposition of Otherwise Impermissible Risks." *Ethical Perspectives* 25 (3): 471–95.
- Sufian, Sandy, and Rosemarie Garland-Thomson. 2021. "The Dark Side of CRISPR." *Scientific American*. <https://www.scientificamerican.com/article/the-dark-side-of-crispr/> (accessed March 20, 2022).



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>