Special Collection: Dual Use, Misuse and Research Ethics

Research Ethics
2025, Vol. 21(2) 245–266

DOI: 10.1177/17470161241272760
journals.sagepub.com/home/rea
Sage

Agent regret and the moral responsibility for the misuse of research results

Jan-Hendrik Heinrichs and Serap Ergin Aslan
Forschungszentrum Jülich, Germany

Abstract

An increasing number of research fields must expect that their projects will be classified as susceptible to misuse or otherwise security relevant, even if the reasons or criteria for this classification have not yet been uniformly developed. Research institutions will commonly distribute the obligation to predict and prevent misuse across multiple members and structures including ethics committees. However, cases of misuse occur even in spite of these precautions, raising the question of the type and distribution of responsibility for the resulting harm. This question becomes even more pressing if research ethics committees in their decisions ask researchers to provide a misuse- specific risk assessment or risk mitigation plan and thereby shift back a part of their responsibility on the researchers. While this kind of requirement may be considered as unsatisfactory by researchers, members of research ethics committees or boards and third-party funders, will argue that it fulfils two important functions. On the one hand, it makes use of the best available source for misuse risk-assessment available to most committees; on the other hand, it guarantees that the researchers at least try to take responsibility. If researchers, committee members and others involved discharge their obligation to predict and mitigate misuse risks with due diligence, any backward-looking responsibility they retain is to show agent regret for the results of their work having been misused.

Keywords

Misuse of research results, responsibility, moral luck, agent regret

Corresponding author:
Jan-Hendrik Heinrichs, Institute for Neuroscience and Medicine 7: Brain and Behaviour, Forschungszentrum Jülich, Jülich 52425, Germany.
Email: j.heinrichs@fz-juelich.de

Introduction

Dual use and misuse are two of the most pressing contemporary issues in research ethics, and for several reasons. Amongst these reasons are scientific developments with high potential for misuse, such as *gain of function* experiments (Imai et al., 2012: 11), political reasons, such as the recent change in the global security situation,[1] as well as the increased attention on prior misuses of science in establishing and sustaining problematic social structures such as the norm of heterosexuality (De Block and Adriaens, 2013) or even slavery (Bynum, 2000).

We take 'dual use' to refer to a research result's potential to be used for purposes that affect the internal or external security of a society. This takes up common distinctions of use such as military/non-military, good/harmful, offensive/protective (Miller and Selgelid, 2008). The reference to internal and external security closely resembles the fields-of-application-approach suggested by Mahfoud et al. (2018), who mention 'political applications [. . .] to govern or manage the conduct of individuals, groups, or populations', 'Security Applications', 'Intelligence Applications' and 'Military Applications'.

Whether a particular use of research results is a case of misuse can be evaluated from an internal and an external perspective. From an internal perspective, it refers to types of use that researchers and their research institutions do not endorse. Given the diversity of researcher's ethical beliefs and research institutions' statutes, this standard can subsume very distinct activities under 'misuse'. From an external perspective, 'misuse' refers to the application of research results to immoral ends. This is not the place to provide detailed a discussion of which ends are immoral. There are, however, some widely shared convictions which predominantly refer to dual use cases as defined above, such as the development of weapons of mass destruction or instruments of torture, and to some widely shared international agreements such as that on biological weapons (Crowley and Dando, 2022). The cases that are of interest to us fall under both the internal and the external standards; that is, they are applications of research results for immoral purposes which are not endorsed by the researchers and their institutions.

Given that we focus on the misuse of research results defined as infringing internal and external moral standards, we presume that the particular activity is (a) morally problematic and (b) judged to be so by the researchers and their institutions. Based on this presupposition we take it to be morally preferable by internal as well as external standards if misuse could be prevented. This rather loose relation of moral preferability has been turned into explicit duties for researchers in several international and national guidelines. Take as an example the World Health Organisation's (WHO) 2022 document: The Global guidance framework for the responsible use of the life sciences: mitigating biorisks and governing dual-use research, which claims: 'Scientists are incentivized to consider, articulate and defend

the potential benefits of their work. They also have a responsibility to consider and mitigate any risks that the knowledge, information, methods, products or technologies that they develop and disseminate could be used for harmful purposes'. (World Health Organisation, 2022: 57)

Similar guidelines are available in several countries, such as the declaration of the German Research Foundation and the National Academy of Science, Leopoldina, which insists: 'Within this complex matrix of benefits and risks, the role of science is to carry out research for the welfare of humankind and the protection of the environment and other values – especially those that are constitutionally protected. Researchers must therefore prevent or minimise direct or indirect harm to values deserving of protection as far as possible'. (Nationale Akademie der Wissenschaften Leopoldina and Deutsche Forschungsgemeinschaft, 2022: 34) Many guidelines have been updated in recent years to react to the increase of potential dual use results in the sciences and their growing potential impact. These revisions often comprised elaborations of the limits, scope and extent of the responsibility of researchers and scientific institutions, referring to principles of non-maleficence or harm-prevention in outlining the duties of scientists in their respective area of application.

What has yet to find sufficient attention in the scientific ethos is the assignment and discharge of responsibilities for misuse of research results which occurs despite these precautions (Ehni, 2008). This omission is relevant – or so we are going to argue – because researchers and their institutions retain a special form of moral responsibility for misuse of research results. This kind of responsibility is not only shown by prominent researchers whose results have been used counter to their moral convictions; it reflects an expectation of lay audiences that scientists – as influential members of society – develop foresight and diligence in their work.

The following will focus on two actors involved in a research project – researchers and ethics board members – and mention a third – leadership, in passing. The reason for this focus is their direct involvement in the deliberation about whether to carry out a specific research project, taking into account its potential for being misused in the sense mentioned above. Admittedly many more actors have an influence on and responsibilities for managing potential misuse as detailed in WHO Global guidance framework mentioned above (World Health Organisation, 2022: xxiii): 'scientists, the scientific community, ethics committee members, institutional and repository managers, biosafety officers, funding bodies, publishers, editors, security officials, regulators, institutional and other authorities, civil society networks, the private sector, other relevant organizations and publics'. Most of these actors are, however, not involved in the decision-making on a specific research project but in designing the social, institutional and physical structures in which such decisions are being carried

out. Where actors are directly involved – for example, biosafety officers – the argument below applies to them as well.

The types of responsibility for research results

Why not focus on researchers' responsibility only? In the end, it is researchers who design and conduct scientific projects and who communicate the results. We will argue that involving additional members of research institutions in the process of scientific research puts a type of responsibility for its results on them, which remains even if they discharged all their duties perfectly. To get this claim off the ground, we first have to make the case that scientific research is a case of collective action, the involvement in which generates a kind of collective responsibility. Otherwise, it would seem implausible that other members of a research institution incur responsibilities for the research's consequences.

Scientific research is predominantly a form of collective, institutional action. Research takes place in government-supported and private organisations such as companies, universities or civil and military research institutes. The times are gone where independent scholars played a discernible role in science. Even – or especially – public involvement in science via participatory designs such as citizen science are closely enmeshed in institutional settings. That does not mean that all of these participants in the process of science have the same institutional, much less the same professional ethical resources at their disposal. They are, however, embedded in a network of interdependent actors.

Accordingly, scientific action is in most cases not *just* individual action – although it is that, too. Rather, scientific actions are – following Miller (2001) – predominantly joint institutional actions. '[J]oint actions are actions involving a number of agents performing interdependent actions in order to realise some common goal' (p. 53), and joint institutional actions are such joint actions, in which the agents are members of one institution and the collective goal is endorsed by that institution. In complex institutions – to which nearly all scientific institutions belong – the members as well as their actions are layered, that is, belong to several different organisational levels – be it in terms of institutional hierarchy or in terms of stratified competences and responsibilities for action. This conception of joint action is compatible with a purely individualist conception of action, according to which collective action is always in principle reducible to individual actions and their arrangement (Bratman, 2014).

Given that joint and institutional *action* consists of layers of interdependent individual action, comparably joint and institutional *responsibility* consists of layers of different types of interdependent individual responsibility. 'On this view of collective responsibility as joint responsibility, collective responsibility is ascribed

to individual human beings only, albeit jointly' (Miller, 2018: 47). Accordingly, the moral responsibility for scientific results is equally joint and thus in parts institutional responsibility.

Members of jointly acting groups and institutions have responsibilities on the basis of their contribution to joint action as well as on the basis of their respective roles and tasks. The latter form of responsibility is based on professional abilities on the one hand; and on the other, simply assigned for the purpose of preventing institutional misconduct or that of taking over liability in institutional spheres of action. As in the analysis of collective action, it is a topic of debate whether collective responsibility is reducible to individual responsibility: Some, such as Pettit (2007), claim that collectives can be responsible for a joint action without any of its members being responsible for it. For current purposes we side with the ontologically more austere position of Miller (2007, 2009) according to which collective responsibility can ultimately be analysed as aggregates of individual responsibility.

What is more, 'responsibility' is not a unified concept. Rather, it is a cluster concept combining a diverse set of normative (and some causal) relations. Thus, the exact normative demands addressed by individual researchers and their institutions under the heading of responsibility need some further elaboration. One version of responsibility, which we claim is of particular importance in cases of misuse, has yet to find sufficient attention in literature.

'Responsibility' has at least four descriptive and – depending on the position in the debate – five or six normative meanings. The following will shortly introduce the most discussed normative versions. However, it should be noted that the prototypical descriptive meaning is that of a cause: Rain is responsible for the road being wet. This is a common way to speak and is presupposed in many (though not all) normative judgements of responsibility (on the following cf. Heinrichs, 2022). The normative meanings can be rendered more precise with the terms: attributability, answerability, accountability, praise- and blameworthiness, liability, obligation and virtue (van de Poel, 2011; Vincent, 2011).

The first five terms refer to the retrospective dimension of responsibility, while the last two refer to its prospective dimensions. In the debate about the so-called faces of responsibility it has become common to differentiate the backward-looking versions at least into attributability, answerability, and accountability (Shoemaker, 2011; Watson, 1996). We will follow this use here and subsume moral liability under accountability as well as assume that there are three different forms of blame and praise, namely attributability-blame (aretaic evaluation), answerability-blame (critique of reasons) and accountability-blame.

Responsibility as attributability (Watson, 1996) refers to the aretaic evaluability of a person on the basis of their behaviour. Responsibility as answerability (Scanlon, 2008) refers to the possibility of critique of a person's reasons and

responsibility as *accountability* refers to a person being subject to reactive attitudes on the basis of their actions (Shoemaker, 2011). The latter is in other contributions to the debate subsumed under responsibility as *liability*, insofar as the reactive attitudes and blame seem to play in morality the role that sanctions take in legal contexts.

An individual's responsibilities can also refer to their *obligations*. Obligations regularly arise from past actions, ranging from wrongdoing to promises and the assumption of social roles. However, there can also be exceptions where responsibility does not depend on prior action. This includes moral duties like the duty to be charitable or provide aid.

Last but not least, 'responsibility' can refer to a *virtue*. Calling someone a responsible individual without further qualification is a common way to praise individuals with certain character traits and predispositions. Responsibility as a virtue appears to be the characteristic that individuals acquire when they regularly fulfil their responsibilities in the meaning variants discussed earlier.

Responsibility for research conduct and research results is thus a form of joint responsibility, that is an interdependent set of different moral relations in which individual researchers, committee members, administrative staff and management are enmeshed within a research institution and beyond. The exact structure of this network will vary not only with each individuals' contribution to the research project in question, but also with the institutional structure and individual obligations within the institution.

On discharging one's responsibilities

Contemporary research environments put the different forms of responsibility for preparing and conducting research projects onto many shoulders. Research institutions have thus successfully addressed one core issue of this type of interdependent and widely distributed responsibilities: the problem of many hands (cf. van de Poel, 2015).

They addressed it – or at least the epistemological and practical dimension thereof – by generating a network of joint responsibility, installing specialised bodies for different tasks, duties and obligations. Research and funding organisations have installed institutional proceedings and structures that help them and their members to shoulder their diverse forms of responsibility. Depending on the infrastructure of an institution, there often are sometimes several ways to get support in preparing research projects: from specialised departments that help with grant applications and shoulder legal liabilities to ethics committees that help with moral decision-making and take over their part in the prevention of becoming morally blameworthy. These departments typically carry different forms of

responsibility for the project, be it for the correct settlement of finances with the funder or the research design following the guidelines in force at the institution.

However, even if every part of this network of agents and responsibilities works as intended (which is admittedly not always the case), there are some occurrences which this network cannot sufficiently cope with, namely events not predictable by and not under the control of the institution and its members. Sometimes research results get misused even though all involved parties acted to the best of their knowledge. Such events out of control of an agent – in the case of institutional agency, out of control of the joint action of the institution's members – raise special challenges for the attribution of responsibility. They are a case of what has been dubbed bad moral luck. In the following we will elaborate this thesis, that misuse of research results beyond the research institution's prediction and control is a case of bad moral luck, insofar as it affects the moral evaluation of the involved parties.

The original problem of moral luck referred to the intuition in everyday morality that agents cannot be made responsible for events not under their control (Nagel, 1976). This so-called *control principle* claims that most normative forms of responsibility can only be ascribed for actions under the control of the agent.[2] It usually refers explicitly to an agent's control over the *action* for which responsibility is ascribed. Thus, the control principle is predominantly a precaution against ascribing agents responsibility for events which they were either coerced into or which merely occurred to them – like being suddenly shoved and consequently bumping into their neighbour. The principle can be extended to cover the *consequences* too, but it thereby loses plausibility. Even the most trivial actions under control of the agent can for unpredictable reasons have excessive consequences. Assigning responsibility, especially in the form of blameworthiness, for such consequences would amount to significant strain on the concept. Therefore, we take the control principle to claim that agents – including institutional agents – are to *blame* only for *events* under their control. Unpredictable consequences of an action are not under the agents control and therefore no reason for blame. That is, we accept that at least as far as *blame* goes, there is no such thing as moral luck, that is no dependence of an agent's blameworthiness on mere luck.

However, even if neither a collective nor an individual agent's *blameworthiness* depends on luck, this does not imply complete irrelevance of luck for the responsibility *tout court* of these agents. Blameworthiness is just one kind of responsibility as explained above. Recent contributions to the moral luck debate differentiated different types of responsibility, insisting that the no-moral-luck intuition is valid for blameworthiness (Sand, 2021) while other aspects of the moral evaluation of an agent including some forms of responsibility might depend on luck.

We cannot elaborate the details of which specific forms of responsibility introduced above does depend on luck in this article. We merely want to claim – with Williams (1976) – that moral evaluation need not be closed when all obligations have been

fulfilled and blameworthiness excluded. Rather we want to claim that as in other cases of bad moral luck such as Williams faultless lorry driver example,[3] there is one specific type of responsibility remaining for the research institution and the involved members and bodies: agent regret.

The term agent regret has been introduced by Williams (1976) to identify the reaction to situations in which an agent faultlessly harmed someone else. It is a feeling that is not directed at either a fault on the side of the agent 'I regret having made a mistake' nor at the harm done to someone else 'I regret the harm done to you', but at the agent having harmed that person. Being faultless is a central criterion for agent regret, because it distinguishes agent regret from guilt and accountability in the form of blameworthiness.[4] That is, the agent recognises that the harm was a bad thing to happen and that they brought it about. The expectation to show agent regret is a form of moral evaluation insofar as we typically think less of a person who, after faultlessly harming someone else, shows no such regret. Thus, this type of moral evaluation in contrast to for example, blameworthiness does in fact depend on luck.

Agent regret takes a peculiar position in the arrangements of forms of responsibility as described above. First of all, it is a feeling and not an action such as paying liabilities or offering excuses. Nevertheless, it is a reaction to previous action and its results, which is expected – if not demanded – of an agent. If an agent faultlessly harmed someone else and neither felt nor showed agent regret, they would provide reason for moral indignation.

Agent regret is quite unlike accountability or answerability in that it can be expected even where a person is clearly faultless, and where the action and its consequences do not reveal anything about the agent's reasons and decisions. It is more like attributability, in that we would attribute a moral deficiency to the person because by not showing agent regret they do not distance themselves from the unwanted effects of an action.

As Vargas (2018) points out regimes of responsibility are justified by their effects. Responsibility in every particular case is assigned following the rules of the regime of responsibility, irrespective of whether the particular case really realises the results by which the regime as a whole is justified. Vargas calls this the 'agency cultivation model', because according to it, the point of a regime of responsibility is to cultivate agency.

Taken together Vargas's agency cultivation model and the differentiation into faces of responsibility as introduced above, we assume that different regimes can regulate different faces of responsibility such as attributability, accountability and agent regret. The specific regime will be justified by its effect on agency, particular ascriptions of responsibility within a regime will be justified by the regime's rules. The rules of the regimes of different regimes of responsibility can well differ, it is for example reasonable to expect that a regime regulating accountability

is different from one regulating attributability (Shoemaker, 2011; Watson, 1996) or agent regret.

The responsibility regime that regulates agent regret is justified by its increasing agents' foresight and diligence concerning possible harm of their actions. As with other forms of responsibility, the justification of a specific case of agent regret differs from the justification of the regime of the specific kind of responsibility.

Thus, individual cases of agent regret need not be rationally justified by their individual effects (MacKenzie, 2017; Rosebury, 1995). Nevertheless, lack of this reaction can be inappropriate and feeling agent regret 'can be appropriate even in cases when questions about one's blamelessness and moral values are not' (MacKenzie, 2017). While its being a feeling distinguishes agent regret from other forms of responsibility, its being appropriate and expected as a reaction to prior, harmful action seems to rank it as a form of moral responsibility.

We take agent regret to be the appropriate and expected reaction of a researcher to the misuse of their research results. The latter seems to be a typical case of harm resulting from a faultless prior action. What is more, to regret having contributed to or made possible the harm incurred by others through information or technology that one's research has generated, seems intuitively to be the correct reaction. If a scientist who explicitly does not endorse some use of their research results – say weaponisation – did not feel any regret for harms incurred by such a weaponisation, we would think their moral reaction to be lacking. In fact, we will later on claim that the same form of regret can be expected from members of research ethics committees.

There are two kinds of evidence for agent regret being expected from and really being felt by researchers. The positive evidence is to be found in behaviour and utterances of scientists, who did indeed regret the use of their inventions. Here are three examples. The first one is a quote from the inventor of nitroglycerin Sobrero, who said: 'When I think of all the victims claimed by explosions with nitroglycerin, and of the terrible devastation they have caused and likely will cause in the future, I am almost ashamed to have claimed recognition as its discoverer. Only the following consideration can console me: sooner or later, nitroglycerin had to be discovered as an inevitable result of chemical research. If not by me, then by another chemist . . .' (Fant, 1995: 133, translation JHH).

The other is that of the inventor of dynamite Alfred Nobel, who intended for the substance to be used for construction and mining purposes and who privately as well as publicly denounced its use in war and terrorism, not least by funding the Nobel Peace Prize. Even if there is an ongoing debate about Nobel's attitude towards the use of his inventions, he is at least by others, including by Albert Einstein in a speech at a Nobel laureate banquet, interpreted to have had and acted upon regrets. Einstein is quoted as 'Alfred Nobel invented an explosive more powerful than any then known — an exceedingly effective means of destruction. To

atone for this "accomplishment" and to relieve his conscience, he instituted his award for the promotion of peace' (Nordlinger, 2012: 18).

Maybe the most impressive examples is Leo Szilard, who conceived the nuclear chain reaction on which the nuclear bomb is based. Szilard did work in the Manhattan project and can be considered one of the main inventors of the bomb. Szilard not only formulated and submitted the Szilard petition asking to demonstrate the bomb only, protested its use in cities, but later in his life wrote a short fiction about his own trial for crimes against humanity and together with Albert Einstein founded the Emergency Committee of Atomic Scientists. For Szilard the use of the nuclear bomb has been a case of misuse and he quite clearly regretted it.

The negative evidence is to be found in writings in the history of science, especially in biographies of famous scientists. There is an expectation in lay and professional audiences to hear about researchers' thoughts on the use of their inventions. Biographies of Nobel, Oppenheimer, Szilard, Haber and others discuss their attitudes towards the uses and misuses of their research because it plays an important role for the appraisal of their character.

Given that research has – as described above – come to be a collective enterprise, we should wonder how far the expectation of agent regret extends. Scientists are already distant causes of the harm that requires any kind of moral responsibility. Neither did Nobel drop a single bomb using his dynamite, nor did he fire a cannon using his ballistite; Szilard also had no hand in the deployment of the nuclear bomb. If they are expected to show agent regret, how much further does this expectation reach? We would like to suggest that the answer to this question depends on two factors: (a) whether the agent was involved with the misused project specifically or had only tangential influence and (b) how remote the influence was. We don't want to go into detail into the remoteness of the agent's influence. It has been sufficiently discussed in the literature about responsibility, how distal causal influence can give rise to responsibility (cf. Sartorio, 2007). However, there might still be an additional difference between two agential influences with the same proximity; namely, whether they were involved specifically with the project in question or had it as the intentional object of their decisions.

Here is an example: Arthur is a chemist who has always taken care to focus his work on purely civil applications. He has studied the relevant guidelines on dual use governance, educated himself on the misuse potential of research results in his field, even thought up and discussed with colleagues what type of results in their field might be misused by bad actors. Nevertheless, one day a method in preparative chemistry he invented is used in a large scale terrorist attack with an otherwise hard to handle poison.

Now imagine we saw Arthur being interviewed about his work after that attack. Imagine he simply insisted that he always demonstrated due diligence, adhered to all relevant rules and guidelines and therefore is beyond reproach and sees no

reason for regret. We – and his colleagues – would think less of him than if he said that he does regret that *his* work has been misused even though he did everything possible to prevent such an event.

A slight modification to the example presents Barbara, a chemist who sat as an expert in the Research Ethics Committee evaluating Arthur's preparative chemistry project. Like Arthur, she fulfilled all her obligations admirably and was justified – given the knowledge available at that time – in supporting the committee's positive opinion of the research project. We conjecture that the same kind of approval or disapproval – if mitigated by degree of proximity – would affect Barbara depending on whether after the attack she showed regret for her involvement in the project's proceeding; that is, agent regret.

To make a contrast: if we place into the same example Charles, who set up the governance structure of Arthur's and Barbara's research institute, we want to suggest that he is not expected to show agent regret. The distinguishing reason is that he was not involved in the decisions and implementation of the research project in question. If he did have an effect, it was through general, non-project-specific decisions and actions.[5] Charles never had Arthur's project as the intentional object of his decisions. Here is a rough analogy to the original literature: we wouldn't think less of the city council setting up traffic rules or the construction workers having built the street if they didn't show agent regret for the accident of Williams' lorry driver.

One could argue against our comparison of Williams' faultless lorry driver to researchers, whose results got misused, that the latter rarely if ever encounter brute bad luck. Rather, researchers often have the option to forgo the benefits of a research project because of its risks – that is discontinue the project – or the option of accepting its risk because of its potential benefits. Whatever decision is made, some likely outcomes are regrettable, but the decision is not faultless.[6]

Against this counterargument we would like to point out that researchers and their peers do not simply make the choice between two risk-benefit ratios but can and must actively manage the remaining risks. Researchers should and do not simply carry on with a security relevant project and accept the risks as they stand. They and their institutions should and often do everything to reduce and mitigate these risks, thereby discharging their forward-looking responsibilities in a way that makes the remaining risks look much more similar to the unpredictable negative outcomes of the faultless driver cases.

The more general claim that a researcher should feel agent regret for the misuse of their results could be rejected by three lines of argument. The first argument simply rejects that there can be a normative requirement to have a specific type of feeling or emotion. Much hinges on the exact formulation here. If the claim is that there cannot be an *obligation* to feel agent regret, this argument is surely correct, but it is not a counterargument to our claim. Obligation is one – forward

looking – type of responsibility, but not the type we discuss here. In particular, it is not a kind of enforceable obligation. We discuss the moral evaluation of individuals on the basis of harms that result indirectly from their actions, which, according to our reading, give rise to a different form of responsibility. The claim – common in the agent regret literature – is that such individuals are judged worse if they do not show agent regret. We take this difference in judgement to imply that such individuals are *expected to show* agent regret, and that this feeling is *judged to be appropriate* on a moral basis. While both are much weaker forms of normative requirement than full blown obligation, it seems barely controversial that such a requirement can extend to emotional reactions.

The second argument claims that the researcher is not involved in the harm caused by the misuse of their research results and thus the criterion of agent regret of faultlessly *harming* someone is not fulfilled. It is not as though researchers instigated malevolent actors to use their research, thus – so the argument – they are not involved in the harm-causing action, thus not harming anyone. If that were correct, it would invalidate the requirement for agent regret. Against this argument, it can be pointed out with Ehni (2008) that according to common criteria there is a weak form of complicity on the side of researchers, if the researcher could have known that there is relevant potential for misuse. Ehni's (2008) colourful comparison is that of A stating, 'This is a stone which can be found in this place and it could be thrown through windows' when there are 'people are around who are searching for stones to throw through windows and the scientist should have known that' (p. 150). Thus, even if the researcher's involvement in the misuse is minimal, it is still sufficient to make agent regret an appropriate reaction. Neither do we think we err on the opposite side of ascribing involvement, which would mean not being able to distinguish involved parties from anyone with even the remotest influence on the events. It is still possible to distinguish those involved in the specific research project from agents who merely had an influence on it by designing the structures in which the project is embedded. Thus, as mentioned above, we take the members of ethics boards who formulated an opinion on the specific project to be under the expectation for agent regret, but not for example the board of directors for setting up the ethics board.

The third argument against agent regret claims that agent regret is not enough, because the harm caused by research misuse does not leave the researcher faultless. Admittedly, if there is a fault on the side of the researcher whose results are misused, this would exclude agent regret as an adequate reaction and place some stronger form of responsibility on them. The above use of the word 'complicity' might be read to insinuate such a fault. However, for the remainder of the article we assume that the researcher in question neither acted negligently, nor were they complicit in any stronger form than unknowingly generating and not explicitly hiding means which others could misuse for their ends.

Given that researchers firstly carry serious forms of responsibility for the misuse of research results if they could have known and prevented or mitigated the risk (World Health Organisation, 2022: 61), and secondly can even be expected to show agent regret when they acted with due diligence, how can the installation and involvement of an ethics board affect this form of responsibility? In a nutshell, we want to argue that the problem of misuse is special in this regard. For nearly all other areas of research misconduct, research ethics committees can help the researchers to discharge their responsibility by providing information and guidance, as well as take over some of the responsibility by guaranteeing that the institution as a whole has paid due diligence. For misuse, however, things look different.

Before we proceed with this argument, here is a quick clarification. Ethics boards or committees have widely differing mandates across different research institutions across the world. Those which do not engage with the misuse potential of research results are not the target of our argument. We exclusively focus on committees which have – usually amongst others – the task and mandate to investigate potential for misuse as defined above, such as the German Commissions for Ethics of Security-relevant research, US Institutional Biosafety Commission or the Ethics Panels in the European Commission's Horizon Europe programme (the last of which do discuss misuse, if not dual use).

The typical input for the proceedings of an ethics board is an application that includes a description of the research project and a self-assessment by the researchers on how they think their project relates to the relevant ethical rules and guidelines. Do they, for example, have a fair participant selection procedure or adequate lab safety standards in place? The typical output of a research ethics institution, whether it is an institutional review board or a research ethics commission, is what has come to be called an opinion. These opinions or recommendations for the main part consist of two analytically distinct parts, namely (a) consent or rejection of an application and (b) terms or requirements. The latter part, the terms, can come in different degrees of obligation – from mere recommendations to strict requirements – and is usually grounded in concrete guidelines or principles. Typical terms make information and competences of the board available to the researchers by advising them that and how they have to draft their consent procedures according to legal and institutional requirements, to adhere to lab safety standards, or to take measures preventing individual profiling in the analysis of online data. Thereby they help to prevent the researcher from breaching specific established norms.

The reaction of a commission or board to a perceived potential for misuse has – for the major part – to follow another logic. Unlike most other terms, it is not a reaction to some possible breach of a research ethical norm by the researcher and their institution. It is not as if the researchers failed to adhere to lab safety standards or to make clear how their research might at least have potential social value. Rather, it is a reaction to potential appropriation of research results by *others* for

purposes which neither researcher nor their institution endorses (Miller and Selgelid, 2007). In consequence, the ethics board cannot request that the research process be adapted to a given and clearly delineated norm.

On the other hand, the suspicion of some project being susceptible to misuse really enables the board to *provide* information, but triggers two informational *demands* on the side of the board: (1) information whether the project is in fact prone to misuse. It might well be the case that the initial suspicion is unfounded for some reason having to do with the project results themselves, with the alleged potential bad actors, with the timing of the results etc. If it turns out that there really is such a risk, (2) information on possible means to reduce misuse risks or misuse impact becomes critical.

As we will discuss in more detail below, both pieces of information are not sufficiently available on the side of the board. This is a fundamental difference to research internal risks, because the board typically is the best place to go to for information on ethical risks of research projects. Members of ethics boards typically have – or at least should have – superior knowledge of local, national and international guidelines as well as competence in research ethics. In the particular case of misuse, however, the best place to go to for information is – if we ignore the security services for a moment– the individual researcher. Thus, the research ethics committees often will have to shift back some burden of providing information, guidance and expertise onto the researchers themselves. This can – and in many cases of, for example, Ethics panels for Horizon 2020 and Horizon Europe funding or Commission for Ethics in Security-relevant Research (KEF) – does take the form of requesting the researcher to provide an assessment of the risks of misuse, and a plan for possible means of mitigation.

This assessment and mitigation plan markedly differs from the abovementioned initial researcher's ethics self-assessment. Initial self-assessment gets evaluated in the committee's proceedings and where necessary corrected. Misuse risk assessment and mitigation plans, by contrast, usually are requested in the evaluation process as an ethics requirement, and even where they are – anticipatory – provided with the initial application they are often not subject to correction by the ethics board, but guide the further proceeding of a research project.

This shift of burden can be perceived as unsatisfactory by researchers, by the leadership of research and funding institution as well as by the committees themselves.[7] Requests for assessments and risk mitigation measures by the researchers have an air of the ethics board shirking its duties. What is more, it seems to put a major part of the burden back on those who have come for guidance, thus sidestepping to shoulder its part of the collective institutional action and responsibility. One might suspect that committees evade even the expectation of agent regret for misuse if they shift so much of the ethical evaluation and the design of precautions back on the researchers. Researchers and the leadership of research institutions could reasonably question the justification for having an ethics board for these cases, and

members of these boards have reason to be dissatisfied with their lack of ability to provide the degree of guidance that is common in other issues of research ethics.

Why researchers should play the key role in discharging of responsibilities

Nevertheless, we want in particular to justify a commission requesting the researcher to provide an assessment and mitigation plan for misuse risks as the best solution within scientific self-governance. We do not think that this procedure alleviates the requirement for agent regret for ethics board members if results get misused anyway.

Our first reason is predominantly negative. It has repeatedly been pointed out that research ethics committees as forms of scientific self-control have significant deficiencies in handling issues of dual use and misuse of research results. In particular, they seem to lack important information which is necessary to evaluate the relevant risks (Resnik, 2010; Salloch, 2018).

Miller and Selgelid (2007) have provided more details on which information exactly scientists – researchers as well as their peers on ethics boards – are lacking in the context of biosecurity in general and weaponisation of smallpox in particular: 'Neither group of scientists, however, would have particular expertise for assessing (1) the likelihood that bioterrorists would be able to access the smallpox virus to begin with, (2) the likelihood that such terrorists would have the means to apply the mousepox technique to the smallpox virus, (3) the likely existence of the intention such terrorists might have to engineer and use such a virus, or (4) the security (as opposed to public health) impact of an attack involving such a virus'. (p. 54) This list can easily be transferred to other cases of misuse of research results. Scientists typically have – beyond not being familiar with possible malevolent actors at all – no information about which technologies these actors have access to, their capacities, their intentions and the security effects of their possible actions. Moreover, this is not the type of information, which scientists can easily get access to, because most of it is classified information and other potential access paths tend to be dangerous or illegal (or both). Given these restrictions, scientists – researchers and board members – are in a bad position to predict and reduce the risk of misuse.

The information that is available for the purpose of *assessing* the misuse risk relates to the technical details of the research project in question: how difficult is it to access and use by non-experts, how widespread is the expertise necessary for its use, does one need special infrastructure to make use of it? This type of information can reliably be found – *qua* experts in the field – with the researchers themselves. While it is possible that some members of ethics committees have this kind of knowledge, too, and clearly should then make it available, this is not

necessarily the case. This is why we take researchers to be in a better, if often still insufficiently good, position to predict the risk of misuse.

What is more, prevention of misuse is not simply a question of continuing or stopping a research project. In many cases, the way a research project is carried out has significant impact on the probability of misuse. Relevant information in this regard is, for example, whether less easily accessible or clearly traceable substances or infrastructures can be used. While knowledge of this kind is available in the research community and there are experts on the prevention of misuse, the specifics of how such information and expertise can be used to modify their specific project, its research question, methodology, or dissemination to *prevent* misuse is predominantly available to the researchers involved. They are the ones most familiar – sometimes the only ones at all familiar – with the exact methodology and context of a project and have a strong incentive to share such specific information only in the dissemination of the project. Thus it is they, who reliably have access to some of the knowledge for risk *prediction* and *mitigation.*

This should, however, not be a reason to relieve research ethics institutions from their duties to review projects for their potential for dual and misuse. It is just to point out the severe restrictions, which they encounter in this task. Information from security services is typically not available to research ethics institutions, and researchers working in the exact same field as a project under review, are neither sufficiently available as a source – either for reasons of bias, if they are within the same organisation, or for reasons of scientific competition and intellectual property if from outside. Thus, the best available source of information are typically the researchers involved in the project themselves and the request for a risk assessment and mitigation plan is simply a committee's best available option.

Our second argument to justify requiring an assessment and mitigation plan is a positive one. Even if a committee or board doesn't have information suited to predict misuse, it does have the information *that* prevention of misuse – and in some institutions that of dual use – is a joint responsibility in the meaning of 'obligation' (see above). This might initially sound trivial, but in fact is not. It is non-trivial to the same degree as the information that research with human participation needs information and consent or that research with animals needs adherence to the 3Rs is trivial. All this information is readily available in guidelines and codes of conduct, but reality in research institutions shows that this type of knowledge (Vinke et al., 2022) – much less the details of how to draft a consent form or how to adhere to the 3Rs (replacement, reduction and refinement in animal research) – is less widely distributed than one would hope.

Thus, while research ethics committees often cannot provide the information about misuse probability and mitigation options themselves, they can make sure that researchers make use of the information they have available or can access it with acceptable effort. In other words: it is the obligation of the research

committee to make sure that the researchers discharge their obligation to prevent misuse as far as possible. As a consequence, they remain involved in the deliberation on the specific project, and therefore share the collective responsibility for its results.

On the responsibility that remains

Nevertheless, the misuse of research results and the success of researchers' attempts to reduce the possibility of misuse depend on external factors that are beyond their control and that researchers and ethics committees have little or no opportunity to assess. Misuse of research results in spite of due diligence and institutional and individual attempts to reduce the risk thereof are a case of bad moral luck for the researchers and their institution.

As introduced above, we do not want to claim that moral luck can influence the responsibility of a research institution and its members in any of the stronger senses such as liability or blameworthiness. What we want to insist on, however, is that the moral evaluation of a research institution as well as of the members involved in a particular research project's planning and conduct does not end when fault and thereby blameworthiness or liability have been excluded. Rather, the institution's members, as far as they were a part of the network of responsibilities involved in the research project in question, can be expected to regret having been the ones to contribute to results which have been misused. This directly implies that members of a research ethics committee which has reviewed a project later misused – or which knew of the project and had the possibility to involve itself – can appropriately be expected to feel agent regret for the misuse. It is them who accepted a research project which later on has been misused and caused the associated harm. Even if faultless, we would think less of them if they did not regret that the research project in the acceptance of which they have been involved was misused. Insofar as agent regret calls for corresponding action at all, they and their institution as a whole would be expected to take up this individual regret in its internal and external communication, in a form that on the one hand clearly identifies, regrets and distances itself from the misuse and on the other hand recognises its members faultlessness.

Some might object that a responsibility regime of agent regret is pointless. If it indeed existed – as claimed above with the examples of Sobrero, Nobel and Szilard – it would be mere personal sensitivities, and if it does not exist that is no loss for science, because it does not play any relevant functional role. Agent regret is too weak to give rise to any operationalisable requirements, and at the same time too demanding by suggesting appraisal beyond fault.

Against this objection we would like to point out that regimes of moral responsibility in general pursue what Vargas has called agency cultivation. A regime of

agent regret does significantly benefit the practice of science if it increases the foresight and diligence of its participants. Science as one of the most influential and novel risk generating fields of human endeavours profits from a rather strict responsibility regime primary by means of in-fact risk reduction, secondarily by means of its external perception as a highly risk-aware community.

There is little doubt that scientists can be held to such a standard. Again, following Vargas, we can for every regime of responsibility identify individuals whose capacity allows them to participate in certain regimes of responsibility and distinguish whether their capacities for this participation just clear the required threshold or clear it by a wide margin. A person's 'modal profile determines whether the agent has the ability to recognise and respond to the relevant considerations, and it comes in degrees. Agents whose modal profile just crosses the required threshold for culpability are distinct from those whose profiles cross the requisite threshold by a considerable margin' (Vargas, 2018: 121). It would seem that scientists are capable of participating in such a regime and clear the regime's capacity threshold by a wide margin. This is why it seems justified to hold scientist, to a fairly strict standard of expectations, namely, to expect agent regret where the harm resulting from their actions cannot be considered their fault at all.

One might raise the opposite objection that by limiting the responsibility for the misuse of research results to a moral expectation to have and communicate agent regret, we excessively exculpate researchers and committee members. However, describing the misuse of research results as a case of luck does not discharge stakeholders from their previous responsibility to reduce the probability of misuse. Rather, it explicitly points to the importance of due diligence in the preparation of research projects. Responsibility is limited to agent regret if and only if all involved parties fully discharged their obligation to identify and mitigate the risks of misuse. For a researcher, the difference between responsibility for misuse in the forms of accountability and mere agent regret lies in assessing the risk and adopting all suitable means for the reduction of misuse – from modifications to the project to its cessation. For the board or committee, the same difference lies in making sure that all available information in the institution is used for misuse risk assessment and mitigation. This is the case irrespective of the fact that a major part of this information usually isn't located with the committee but with the individual researchers.

We think that the obligation to identify and mitigate the risk of misuse implies a further obligation for committee members or other, specialised members of the institution to keep informed about possible dangers of misuse and common means to mitigate them. This is strictly parallel to a researcher's duty to keep themselves informed about the standards of Good Scientific Practice as included in many national and international guidelines as well as the duty to prevent risks from research as formulated in several institutional standards. Presently there are significant hurdles in fulfilling such a duty, because the information on prior cases of

misuse are often classified and information on the applications and reviews of research ethics committees are – for reasons of IP – often confidential. Nevertheless, keeping informed via the available sources seems to be required of committee members. While there might be different resources available in some cases, a specific project's researchers will often be the best source. Thus, a committee which does not urge researchers to use and provide their knowledge for misuse risk mitigation will – at least morally speaking – incur a greater responsibility than mere agent regret.

Acknowledgements

We would like to thank all the participants of the project DUMFE: Dual Use and Misuse of Research Result, its funder the German ministry for education and research (BMBF), the members of the Commission for Ethics in Research and of the Working Group Neuroethics and Ethics of Artificial Intelligence of Forschungszentrum Jülich

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

All articles in Research Ethics are published as open access. There are no submission charges and no Article Processing Charges as these are fully funded by institutions through Knowledge Unlatched, resulting in no direct charge to authors. For more information about Knowledge Unlatched please see here: http://www.knowledgeunlatched.org.

The project has been funded by the German Ministry für Education and Research (BMBF) Project DUMFE (Dual Use and Misuse of Research Results), 01GP2187

Ethics approval

This research project did not require ethics approval.

ORCID iDs

Jan-Hendrik Heinrichs https://orcid.org/0000-0002-0560-8483
Serap Ergin Aslan https://orcid.org/0000-0003-0341-1000

Notes

1. As succinctly put by the U.N. Secretary General at the opening of the High-Level Segment of the 55th session of the Human Rights Council at the European headquarters of the United Nations in Geneva, Switzerland, Monday, Feb. 26, 2024, who claimed that the world is 'less safe by the day'.
2. The resulting paradox introduced by Nagel is that by assigning responsibility or basing moral evaluation on events under our full control only, we threaten to shrink the scope of moral evaluation to an extensionless point. We cannot go into any details of the debate in the present article, because we only rely on the very limited claim that one type of moral

evaluation which indeed is subject to luck, is the expectation to feel agent regret. For a good starting point into the moral luck debate (beyond the original articles by Nagel and Williams) see Enoch and Marmor (2007), Raz (2012). and for an empirical perspective Kumar (2019).

3. The example is introduced in Williams 1976: 124: 'The lorry driver who, through no fault of his, runs over a child, will feel differently from any spectator, even a spectator next to him in the cab, except perhaps to the extent that the spectator takes on the thought that he might have prevented it, an agent's thought'.
4. However, some such as Blackburn (2015) or Enoch and Marmor (2007) might dispute this distinction, arguing that feeling agent regret implies thinking that one was or could have been at fault.
5. Admittedly there is a vague area to this distinction. Such a borderline case would be Deborah, a journal editor ion one of Arthur's articles. She (a) has a more direct influence on potential misuse and (b) is involved in the dissemination and thus directly in the decision about one part of the project. On the other hand, she comes in after the deliberation about the whether and how to pursue the research project itself.
6. In this, researchers cases are much more like Williams' Gaugin case where one is faced with a dilemma and must choose between two negative results.
7. One of the authors has received feedback from several research consortia as well as members of funding organisations uttering their discontent with risk mitigation plans as an ethics requirement resulting from committee decisions – while such plans actually are a useful part of the original ethics application to such a committee. There is, however, no systematic study on the different agents' attitudes towards this kind of ethics requirement. On the expectations of researchers towards their ethics committees see Keith-Spiegel and Tabachnick (2006).

References

Blackburn S (2015) Williams, Smith, and the peculiarity of piacularity. *Journal of the American Philosophical Association* 1(2): 217–232.

Bratman M (2014) *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.

Bynum B (2000) Discarded diagnoses: Drapetomania. *The Lancet* 356(4): 1615.

Crowley M and Dando M (2022) *Toxin and Bioregulator Weapons: Preventing the Misuse of the Chemical and Life Sciences*. Cham: Palgrave Macmillan.

De Block A and Adriaens PR (2013) Pathologizing sexual deviance: A history. *Journal of Sex Research* 50(3–4): 276–298.

Ehni H-J (2008) Dual use and the ethical responsibility of scientists. *Archivum Immunologiae et Therapiae Experimentalis* 56(3): 147–152.

Enoch D and Marmor A (2007) The case against moral luck. *Law and Philosophy* 26(4): 405–436.

Fant K (1995) *Alfred Nobel: Idealist zwischen Wissenschaft und Wirtschaft* (W. Butt, Trans.). Basel: Birkhäuser. Available at: https://books.google.de/books?id=ttogAQAAIAAJ

Heinrichs J-H (2022) Responsibility assignment won't solve the moral issues of artificial intelligence. *AI and Ethics* 2(4): 727–736.

Imai M, Watanabe T, Hatta M, et al. (2012) Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* 486(7403): 420–428.

Keith-Spiegel P and Tabachnick B (2006) What scientists want from their research ethics committee. *Journal of Empirical Research in Human Research Ethics* 1(1): 67–82.
Kumar V (2019) Empirical vindication of moral luck. *Nous* 53(4): 987–1007.
MacKenzie J (2017) Agent-regret and the social practice of moral luck. *Res Philosophica* 94(1): 95–117.
Mahfoud T, Aicardi C, Datta S, et al. (2018) The limits of dual use. *Issues in Science and Technology* 34(4): 73–78.
Miller S (2001) *Social Action: A Teleological Account*. Cambridge: Cambridge University Press.
Miller S (2007) Against the collective moral autonomy thesis. *Journal of Social Philosophy* 38(3): 389–409.
Miller S (2009) *The Moral Foundations of Social Institutions: A Philosophical Study*. Cambridge: Cambridge University Press.
Miller S (2018) *Dual Use Science and Technology, Ethics and Weapons of Mass Destruction*. Cham: Springer.
Miller S and Selgelid MJ (2007) Ethical and philosophical consideration of the dual-use dilemma in the biological sciences. *Science and Engineering Ethics* 13(4): 523–580.
Miller S and Selgelid MJ (2008) *Ethical and Philosophical Consideration of the Dual-Use Dilemma in the Biological Sciences*. Dordrecht, Netherlands: Springer.
Nagel T (1976) Moral luck. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 136–151.
Nationale Akademie der Wissenschaften Leopoldina and Deutsche Forschungsgemeinschaft (2022) *Scientific Freedom and Scientific Responsibility Recommendations for Handling of Security-Relevant Research*. Halle (Saale). Gemeinsamer Ausschuss zum Umgang mit sicherheitsrelevanter Forschung von DFG und Leopoldina.
Nordlinger J (2012) *Peace, They Say: A History of the Nobel Peace Prize, the Most Famous and Controversial Prize in the World*. New York, NY: Encounter Books.
Pettit P (2007) Responsibility incorporated. *Ethics* 117(2): 171–201.
Raz J (2012) Agency and luck. In: Heuer U and Lang G (eds) *Luck, Value, and Commitment: Themes From the Ethics of Bernard Williams*. Oxford: Oxford University Press, 133–162.
Resnik DB (2010) Dual-use review and the IRB. *Journal of Clinical Research Best Practices* 6(1): 1–6.
Rosebury B (1995) Moral responsibility and "moral luck". *The Philosophical Review* 104(4): 499–524.
Salloch S (2018) The dual use of research ethics committees: Why professional self-governance falls short in preserving biosecurity. *BMC Medical Ethics* 19(1): 53.
Sand M (2021) A defence of the control principle. *Philosophia* 49(2): 765–775.
Sartorio C (2007) Causation and responsibility. *Philosophy Compass* 2(5): 749–765.
Scanlon T (2008) *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge: Belknap Press of Harvard University Press.
Shoemaker D (2011) Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics* 121(3): 602–632.
van de Poel I (2011) The relation between forward-looking and backward-looking responsibility. In: Vincent NA, van de Poel I and van den Hoven J (eds) *Moral Responsibility: Beyond Free Will and Determinism*. Dordrecht, Netherlands: Springer, 37–52.
van de Poel I (2015) The problem of many hands. In: van de Poel I, Royakkers L and Zwart SD (eds) *Moral Responsibility and the Problem of Many Hands.* New York and London: Routledge, 50–92.

Vargas MR (2018) The social constitution of agency and responsibility: Oppression, politics, and moral ecology. In: Oshana M, Hutchison K and Mackenzie C (eds) *Social Dimensions of Moral Responsibility*. Oxford: Oxford University Press, 110–136.

Vincent NA (2011) A structured taxonomy of responsibility concepts. In: Vincent NA, van de Poel I and van den Hoven J (eds) *Moral Responsibility: Beyond Free Will and Determinism*. Dordrecht, Netherlands: Springer, 15–35.

Vinke S, Rais I and Millett P (2022) The dual-use education gap: Awareness and education of life science researchers on nonpathogen-related dual-use research. *Health Security* 20(1): 35–42.

Watson G (1996) Two faces of responsibility. *Philosophical Topics* 24(2): 227–248.

Williams B (1976) Moral luck. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 115–135.

World Health Organisation (2022) Global guidance framework for the responsible use of the life sciences: Mitigating biorisks and governing dual-use research. Available at: https://www.who.int/publications/i/item/9789240056107 (accessed 23 August 2024).