Responsibility: Identifying Purpose and Finding Meaning

Luke Price*

Abstract: The social and legal practices of blaming, praising, punishing and rewarding are inextricably linked with the process of ‘holding responsible’. Blame, praise, and the like exist as means of holding agents to account that is distinct from, but reliant upon, attributions of responsible agency. When claims of accountability are made without access to an underlying shared attribution of responsibility, the communicative role of accountability is undermined. Disagreement over blame and praise is reduced to disparity: able to hear only that something is a bad or good thing, we are left unable to understand what the bad thing is, or why it is bad.

Responsibility as we employ it offers the basis for our evaluations of agents. However, conceptions of responsibility that focus on agent capacities, namely control and rationality, fail to give responsibility a meaning capable of fulfilling this purpose. The retrospective responsibility of agents for events does not result from the capacities of those agents. It is attributed on the basis of agents’ roles in events, enabling accountability for those roles.

Keywords: Responsibility, Blame, Punishment, Agency

* University of Exeter.

1. Introduction

The concept of responsibility defines the conduct of human relationships. We employ responsibility as part of our social and legal practices, in the positive or negative positions we adopt towards other agents. Responsibility offers a means of holding others to account, through the communication of evaluative attitudes of blame and praise. Communicating these attitudes is fundamental to human interaction. However, such communicative practices cannot be based solely upon our relationships with one another.

Accountability attitudes of blame and praise are a product of evaluative stances. They are defined by personal positions and values. If responsibility is to communicate attitudes of accountability, then it must retain its meaning across boundaries of evaluative disagreement. To survive disagreement, evaluative attitudes must be anchored in shared content, or else lose their meaning, and consequently, purpose.

Accountability, in enabling the interaction of evaluation that is fundamental to human relationships, must be grounded in the shared reality that encompasses those relationships. The communication of accountability enables interaction between agents, but is enabled by those agents' interaction with the world: by the relation between agents and events. It is in the need to encompass this relationship in responsibility practices that the role of attribution is found.

Responsibility attributions describe the connection of agents to events. In doing so, they provide access to the shared reality that makes it possible to communicate attitudes across lines of evaluative disagreement. Responsibility attributions accompany accountability, as a necessary pre-requisite of our evaluative attitudes towards others.

This purpose of responsibility has significant consequences for the requirements of its form. Any concept of responsibility must demarcate responsible agency prior to the ascription of blame and praise, by reference to transevaluative conditions of responsibility. If it is to do so, then it cannot depend upon evaluative positions, or lose its meaning when subjected to evaluative disagreement.

Responsibility is often treated as a status attained by agents, and from this defined according to agent characteristics – commonly the possession of control or rationality. These approaches are incapable of satisfying the requirements of a responsibility that enables attitudes to traverse evaluative disagreement. In order to reach a responsibility concept that is functional in fulfilling both its purpose and meaning, there is a need to move away from understanding responsible agency as a matter of agents' characteristics. Responsible agency results not from the capacities agents hold, but from their roles within events. By looking at chains of events and identifying responsible agents within them, accountability attitudes find meaning, and responsibility attributions purpose.

2. Identifying responsibility's purpose within social practice

Responsibility just for its own sake is no responsibility at all. The practice of attributing responsibility always involves more than a simple claim that someone is responsible for something. It always involves an evaluation of an agent on the spectrum of blame and praise. In theory, one could imagine I might be held responsible for moving a pencil on my desk ten centimetres to the left. This would never occur, however, as there is no evaluation of me to be made on the basis of such normatively insignificant action.

Whenever responsibility is employed, it is because there is an evaluation of an agent to be made, such that the agent is held to account for the normative consequences of their behaviour. Responsibility statements are made only in situations in which they are normatively significant. We make attributions of responsibility because they are to be followed by an evaluation of the agent; if they were not then no attribution would be made. Responsibility is inherently normative, and coexists with evaluation.

From responsibility ensues evaluation, in the form of accountability attitudes on the spectrum of blame and praise. Responsibility, as practiced, is always connected with adopting an evaluative position towards agents.¹ However, the normativity of responsibility statements does not imply that they vary with evaluative stances. It is a mistake to assume the possibility of meaningful communication across lines of evaluative disagreement in the absence of transevaluative, mutually understood content. In truth, the communication of evaluative attitudes between disagreeing parties is only possible when anchored by shared transevaluative understanding.

A. Responsibility and evaluative attitudes

When holding someone accountable, we make evaluative claims against them – we express attitudes. However, our evaluative claims are more than positive or negative attitudes, or statements of moral positions. Evaluations are accompanied, and supported, by references to content independent of our evaluative positions. This transevaluative content is essential to the communication of evaluative attitudes.

Attitudes of accountability are central to human interaction. However, these relationships are frequently comprised of evaluative disagreement. By way of example,

¹ P.F. Strawson, *Freedom and Resentment and Other Essays* (Methuen 1974), 16-17.

we hold individuals accountable for being ‘terrorists’, to enable social blame and legal punishment, yet concepts like ‘terrorist’ do not exist unchallenged. Someone I call a terrorist may be another’s freedom fighter. We may perceive and describe exactly the same act, but evaluate it differently.

To hold an agent accountable for terrorism is to adopt an evaluative position towards his behaviour. An evaluative stance on its own, without (implicit) reference elsewhere, does not permit the two notions of ‘terrorist’ and ‘freedom fighter’ to interact. There is nothing they share. It is only with some common frame of reference, allowing us to see what is being evaluated positively and negatively, that the two concepts can interact. We are only then able to delineate the content of our differing positions, such that we can see why our positions differ: not simply that they do differ.

Accountability is concerned with human relationships, but cannot be fully explained by reference to those relationships. The relationship between two positions may be one of evaluative difference, which without anything further offers no means of communication. Accountability attitudes are in this way undermined, prevented from being communicated by their lack of meaning beyond their evaluative point of origin.

In practice, we find that we can have different evaluative positions and still understand each other’s accountability attitudes. This is because we disagree within the context of a shared human reality: an existence in which relationships are formed not only between individuals, but between individuals and the world: between agents and events. This common relation to the world provides the frame of reference for our evaluative disagreements. Human experience is shared, such that whilst we may disagree over terrorism, we do so through concepts that form part of the concept of terrorism, such

as swinging a knife, pushing a button, or pulling a trigger. Though these basic concepts may seem simple, and are easily taken for granted, they enable evaluation to retain its meaning.

We might disagree over whether these things occurred; we might see them as differently motivated and thus different events; or we might evaluate them differently. All of these factors make up part of our contrasting attitudes of accountability. However, our evaluative disparity has expanded into disagreement. It is no longer a matter of two countervailing and irreconcilable positions. We have a reference point from which to understand the other evaluative position, from which to meaningfully disagree. We are able to determine whether we perceive the event differently, apply a different description to what we perceive, or evaluate our perceptions differently. Our countervailing evaluative attitudes become more than mere disparate opinions.

Human experience provides the tools with which we can communicate accountability attitudes in the face of evaluative disagreement. Our shared existence provides not common meaning, but common understanding. Evaluative attitudes are subject to disagreement, rather than disparity. Reference to the interaction between agents and events allows accountability to inhabit its communicative role, as a means of conveying evaluative attitudes.

A common relationship with the world gives meaning to our relationships with one another, making it possible for disagreeing parties to understand what it is that they are disagreeing about. This content exists beyond our evaluative stances, as an intrinsic component of human experience. Something when we communicate blame is providing

access to this transevaluative content, ensuring the understanding of our accountability attitudes. It is here that we find the role of responsibility.

Responsibility attributions persist across evaluative positions, remaining consistent across the spectrum of different attitudes. The praised or pardoned party is no less responsible than the guilty, though the attitudes that follow as consequences of their responsibility differ. Whilst their accountability changes, the accompanying attributions of responsibility do not. The alcoholic who cannot stop drinking and so is the subject of sympathy, and the friend whose lateness we have come to accept, are both still responsible for their behaviour. Our attitudes to them simply differ from those we would have towards a more deliberate drunk or a new acquaintance. There is a crucial difference between 'not to blame' and 'not responsible'. Responsibility does not change with our evaluations.

Attributions of responsibility describe agents' relations to events. Without attribution, accountability attitudes lose their meaning across evaluative stances, and their purpose as a means of facilitating communication across those positions. When accompanied by responsibility attributions, evaluative attitudes can inhabit human relationships, enabling communication by reference to common human experience. Responsibility attributions, through their reference to shared reality and consequent independence from evaluative positions, provide the transevaluative content that makes accountability possible.

Responsibility and attitudes are both coexistent and co-dependent. The expression of attitudes through which we hold others accountable relies upon a transevaluative

attribution of responsibility.² Equally, however, responsibility is only attributed when there are attitudes to be expressed. The attribution of responsibility and expression of attitudes are mutually dependent. Responsibility attributions and accountability's attitudes are distinct, but inseparable components of evaluative communication. There is no responsibility without evaluative attitudes, and no accountability without responsibility attribution.

Responsibility operates as the distinct transevaluative precursor for our evaluative attitudes. Responsibility attributions, as an unmoving counterpart to evaluative attitudes, provide access to the points of reference we rely on to engage in blame or praise. Whenever we blame or praise, we are using transevaluative attributions of responsibility to cross boundaries of evaluative disagreement.

It is difficult to imagine a world without access to the shared transevaluative content of our disagreements. The notion that we could not communicate our shared human experience with others seems fanciful. That this is the case indicates how important responsibility attributions are, and how often they are being used to traverse the gaps between our evaluative positions. A pure evaluation of 'freedom fighter' competing with another of 'terrorist' offers no common concepts to form the basis of our evaluative claims. Both terms necessitate a means of filling the void of content. Responsibility performs this function, all the time, and is of vital importance within our communicative practices.

By providing the agreed-upon content on which our attitudinal statements are implicitly based, responsibility enables attitudes, both negative and positive. Treating someone as responsible, recognising the extent of their responsibility, enables sympathy

² Gary Watson, 'Two Faces of Responsibility' (1996) 24 Philosophical Topics 227, 229.

and praise. It allows us to pardon or forgive as well as blame and punish. In all of these cases, with the benefit of responsibility, we are able to demarcate the boundaries of our descriptive, perceptive, and evaluative differences. We can see the full picture of our disagreement, and understand what attitudinal claims we are making. Without responsibility, we would have no tool with which to access the shared content that enables communication across lines of disagreement.

B. Framing responsibility's form by reference to its function

The transevaluative role of responsibility has significant implications for its necessary form. Responsibility must be determined by reference to common conceptions of agents' relationship to the world, shared across evaluative stances, so as to remain the product of transevaluative content. Disagreements over the evaluations that follow responsibility, and over responsibility itself, are only made possible by reference to accepted basic concepts shared by disputants. Replacing these basic concepts with the products of evaluative positions would, in the absence of metaphysically true and universally accessible moral facts, introduce disparity into the content that must be shared if it is to provide the agreed-upon points of reference for our dispute.

Though evaluative disagreement does not disprove the possibility of metaphysical moral objectivity, it does show that epistemological access to such truths is beyond the reach of disagreeing parties. If such access were possible, then there would be no disagreements to resolve. Moral facts cannot provide the transevaluative basis for evaluative disagreement. Dependence on evaluative positions offers no common content with which to communicate evaluative attitudes. Only reference to the shared experience

of human reality can preserve the possibility of communicating accountability across evaluative boundaries.

There are limits to the reach of shared understandings to which responsibility provides access. Some individuals will not apply these concepts, predominant across human experience though they may be. Thus, an individual may consider breaking a china plate to be a killing and an act of terrorism. That a small percentage of agents employ basic concepts distinct from those shared within the human community does not undermine that community. Some individuals operate outside of our shared experience and accepted understanding. They do not interact with other individuals, and with the world, in the same way. The end result is one that is common to many theories of responsibility: the responsibility of those who equate plate breaking to killing is denied by their mental disorder.

Responsibility reaches only so far as our shared human experience, and there will be those who do not adopt the broadly accepted perceptions of sane individuals.³ They operate outside of the world we collectively inhabit, and consequently we lack the shared frame of reference needed to make sense of their evaluative position. We have lost the mode of communication with them that renders our responsibility attributions meaningful. We punish to deter or to scold, or praise to encourage, but never communicate, as this relies upon the common relationship to the world that exists as a key part of shared human experience.

We rely upon a common reality to make sense of the attitudinal statements we make across lines of evaluative disagreement. Responsibility can provide a means of accessing these shared understandings. It allows access to the content understandable by

³ Isaiah Berlin, *Four Essays on Liberty* (Oxford University Press 1969), 204.

both sides of an evaluative disagreement that permits those parties to understand and engage with each other. It can in this way keep us from talking past each other.⁴

Responsibility is not evaluative, but is a necessary component of our evaluative attitudes.

3. Identifying responsibility that enables attitudes

Just as attitudes of blame and praise depend upon responsibility attributions in order to be understandable, our means of defining and demarcating responsibility must be able to explain the attitudes we adopt and practices we engage in. Responsibility must identify the distinction that is drawn between a domino that topples onto another and a human adult who knocks down a stranger. Equally, a concept of responsibility must not conflict with the distinctions we draw between an adult human's intentional actions and an infant's accidents. Crucially, it must explain why we make those distinctions by making reference to shared transevaluative content, thereby enabling disagreement over, and communication of, evaluative judgments of blameworthiness or guilt.

A. Searching for the capacity to be responsible: control

A common means of demarcating the boundaries of responsibility is to point to the capacities of the agents we are calling responsible. We seek to deny the responsibility of a range of causal agents: whether they are storms, trees, infants, or some non-human animals. These agents seem to lack certain capacities, and hence we point to the capacities absent in the non-responsible as grounds for responsibility.

⁴ Allan Gibbard, *Thinking How to Live* (Harvard University Press 2003), 282.

One such capacity is control. If control is taken as a condition of responsibility, then we are responsible when we are in control of what we do.⁵ An agent's responsibility may be made conditional upon that 'person's ability to control her mental states and events leading to action in the relevant causal way, such that she is able to perform actions freely and thereby cause events to occur.'⁶ On this account, when we lose control of ourselves or never had it in the first place, whether by reason of a magician's hypnosis or physical incapacity, we are not responsible. However, when I choose to knock someone else over, I may be distinguished from a domino that uncontrollably topples into another, and can be deemed responsible.

The control requirement imposes a threshold below which acts will be deemed outside the boundaries of responsibility. The exercise of control, as an element of moral agency, signifies the moral quality of, and agent's responsibility for, the act in question.⁷ It is this moral quality that gives rise to blame and praise.⁸ Whatever form control is required to take, it will always be a test of sufficiency, one that must be satisfied if acts are to be of the quality necessary for responsibility.

The focus of discussion is typically concerned with the manner in which the control requirement should be framed: how control must be exercised in order for an agent to be deemed responsible.⁹ However, prior to the issues arising from the implementation of the control requirement is the challenge presented by the inherent dependence on evaluative judgments that results from determining the responsible quality

⁵ Robert E Lane, 'Moral Blame and Causal Explanation' (2000) 17 *Journal of Applied Philosophy* 45, 55.

⁶ Walter Glannon, *The Mental Basis of Responsibility* (Ashgate 2002), 22.

⁷ Stephen J Morse, 'Culpability and Control' (1994) 142(5) *University of Pennsylvania Law Review* 1587, 1606.

⁸ *ibid.*, 1589.

⁹ See for example John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press 2000).

of acts on the basis of agents' control. By failing to assess responsibility on basis of shared transevaluative content, introducing control as a condition of responsibility undermines the possibility of responsibility being used as a means to render evaluative attitudes understandable.

Crucially, control is never absolute. If I attempt to fire a well-maintained gun and it misfires, then I remain responsible for pulling the trigger. However, my good fortune in this instance means no bullet left the chamber. External factors will almost always affect my acts, changing the outcome for better or worse. Controlled action is conditioned by this moral luck.¹⁰ Two drivers may be equally negligent, but that negligence may have substantially different consequences if only one suffers the misfortune of injuring a pedestrian.¹¹ With so much occurring contemporaneously with our actions, there is no way we can be said to be in absolute control.

We might, as some do, treat equally negligent drivers as equally responsible without regard to the consequences of that negligence.¹² However, whilst it is possible to exclude clear cases of resultant luck, luck can never be excluded entirely from responsibility. Luck changes both the meaning and the consequences of individual acts. We can be highly influential, but our actions, both possible and realised, are conditioned by external contingencies that we can neither predict nor control. Our lives are governed by moral luck and the unpredictability of consequences.¹³ Requiring total control would undermine our responsibility, because we are never in control of the world; the

¹⁰ Bernard Williams, *Moral Luck: Philosophical Papers 1973-1980* (Cambridge University Press 1981), 21.

¹¹ Joseph Raz, *From Normativity to Responsibility* (Oxford University Press 2011), 251.

¹² Thomas Nagel, *Moral Questions* (Cambridge University Press, 1979), 25.

¹³ Hannah Arendt, *The Human Condition* (2nd edn, University of Chicago Press 1998), 233.

unexpected can always happen. If we were only responsible when we were in control, then we would never be responsible at all.¹⁴

Total control cannot be required for responsibility, and the question then arises of the extent of luck, as a constraint on control, that still permits agency and responsibility to persist.¹⁵ If an individual is unfortunate enough to find himself suffering from borderline personality disorder, then we may disagree over whether that agent has control sufficient to imbue his acts with the quality of responsibility, depending on our evaluation of the effect on moral agency of constitutive luck that produces a genetic predisposition to BPD. I may give that agent's condition significant weight, thus deeming them out of control and not responsible, whilst another considers genetic conditions not to undermine an agent's control or the corresponding responsible quality of their acts.

The assessment of control sufficient for responsibility must rely on some reason why that particular degree of control suffices. It does so because a judgment of accountability has been made. Rather than responsibility enabling evaluative attitudes, the expression of evaluative attitudes justifies responsibility. Attributions of responsibility are not made because something is a matter of evaluation, or of normative significance; instead responsibility only attaches when its subject matter is at a particular point on the scale of evaluative attitudes. An agent is responsible for something only in accordance with its particular evaluative weight. Responsibility attributions are made because what has happened has been evaluated. Responsibility is rendered dependent not just on the possibility of evaluation, but on the making of particular evaluations.

¹⁴ Nagel (n 12), 26.

¹⁵ *ibid*, 25.

Using blame as an example, we might find fault with only some degrees of negligence, such that not all negligent drivers are equally responsible. Because responsibility is a condition of the evaluative judgment of fault, only culpably negligent drivers would be deemed responsible. Responsibility is thereby conditioned upon behaviour in relation to which control is deemed to exhibit a moral defect or proficiency, such that in the case of negligent drivers 'we hold people responsible for exercising certain capacities of self-control that we (correctly or incorrectly) assume normal beings to possess.'¹⁶

Conditioning responsibility upon moral evaluations provides means of setting boundaries of how much control is required for responsibility. Culpability is defeated by insufficiency of control. Any act that is beyond the control of the agent such that it does not warrant blame or praise is not the responsibility of that agent. Only the driver blamed for his negligence is responsible, because only he behaved in a substandard way.¹⁷ Perhaps he paid less attention to the road, or the condition of his car, than we reasonably expected him to, and was therefore at fault. As a result, we hold him responsible.

If moral evaluation is given the role of justifying responsibility attributions, it becomes the case that what we are responsible for is dependent upon what we are at fault for. Not just attitudes of accountability, but also attributions of responsibility, become a product of our evaluative stance. Our responsibility, as a consequence, is dependent upon evaluative positions. We lose the common ground between cultures that is essential to a purposeful concept of responsibility. Responsibility ceases to be a transevaluative concept, becoming one that is wholly reliant upon the evaluations made. If we are to

¹⁶ Ronald D Milo, *Immorality* (Princeton University Press 1984), 88.

¹⁷ Julie Tannenbaum, 'Emotional Expressions of Moral Value' (2007) 132 *Philosophical Studies* 43, 51.

preserve responsibility across disagreement, a faulty agent's responsibility should not differ from the perfect agent's. Fault may be relevant to blame, but not to responsibility.

When control is taken as the grounds of responsibility, responsibility is rendered dependent on evaluative judgments. The evaluation of what and how much luck is sufficient to undermine the responsible quality of an agent's act determines whether or not that agent is understood as responsible. Responsibility thereby loses its transevaluative meaning. What results is the impossibility of access to shared human experience with which to make our evaluations. The purpose of responsibility is undermined when it is based upon a judgment of the existence of control.

B. Alternate capacities: responsible when rational?

We might instead try to equate responsible agency with rationality. Thus, when we are rational, we may be said to be responsible.¹⁸ However, if we are responsible, or particularly responsible, when we are rational, then the question of what reasons are right or best will determine our responsibility. Our status as rational is a product of the reasons we are deemed to have had: it is a product of an evaluative judgment, dependent upon an evaluative position.

Responsibility is a means of enabling evaluation in the face of disagreement. It enables speakers to express evaluative attitudes in a way that the listener can understand, even across lines of evaluative disagreement. From responsibility having this function, of conveying the speaker's position in an understandable way, it follows that responsibility is always communicated. If rationality is the basis of responsibility, then when attributions of responsibility are made, claims of rationality are also made.

¹⁸ Vincent Chiao, 'Action and Agency in the Criminal Law' (2009) 15 *Legal Theory* 1, 16.

The reasons that form part of responsibility attributions are both spoken and heard. Regardless of whether reasons justifying responsibility are understood as internal or external to the agent who holds them, they are always externally perceived.¹⁹ Responsibility is not just held, but attributed. What matters for rationality, and therefore responsibility, is the external perception of reasons. The reasons relevant to responsibility are those perceived by the person who is making the responsibility attribution. Consequently, the primacy and adequacy of reasons are viewed from an external perspective. The rationality that determines responsibility will be a product of the values we apply in deciding what to perceive as reasons.

When assessing what reasons someone had, we adopt evaluative stances. Therefore, when we make claims of rationality, we make evaluative claims. If responsibility is rationality, then claims of responsibility are evaluatively dependent. When we disagree using a rationality-based conception of responsibility, our responsibility attributions become subject to the same disparity as our evaluative attitudes. Imposing responsibility for rationality invokes claims of value that responsibility must supervene if it to provide a frame for understanding disagreement.

Responsibility does not resolve disagreements between evaluative claims – it enables disagreement, by creating a context that makes evaluative attitudes understandable. In making evaluative claims we may disagree over matters of evaluation, perception, or description. We may also disagree over whether an agent is responsible. However, evaluative claims are not possible without a shared concept of what responsibility is. A model of rational responsibility creates disagreement not just over the evaluations that accompany responsibility, but over the concept of responsibility itself.

¹⁹ Williams, *Moral Luck: Philosophical Papers 1973-1980* (n 10), 101.

When rationality is relied upon for responsibility, responsibility becomes a product of evaluative perspectives of what is rational, of what reasons an agent had, rather than our shared understanding of the world. We lose the function of responsibility as a tool for making sense of accountability attitudes, as it becomes dependent upon evaluative stances, not shared human experience. Responsibility loses its purpose as a means of enabling communication across evaluative disagreement. With standards of rationality dependent upon evaluative stances, a rationality-based responsibility reduces our blaming and praising attitudes to opposing non-interactive evaluations. Responsibility becomes dependent upon evaluative stances, rather than common human content, and ceases to be able to explain our blaming attitudes or our disagreements over blame.

One possible avenue for avoiding this evaluative dependence is to seek to detach rationality from its connections with the right or wrong course of action. Raz's approach is of this sort: an attempt to tie responsibility not to rational action, but to 'rational functioning', whether it is expressed through rational or irrational action. Raz's position is defined by the Rational Functioning Principle (RFP), under which:

'Conduct for which we are (non-derivatively) responsible is conduct that is the result of the functioning, successful or failed, of our powers of rational agency, provided those powers were not suspended in a way affecting the action.'²⁰

This approach looks to be premised not on rational failure or success, but on the exercise of rational capacities generally. Whether successful or not in the exercise of our rational capacities, 'we are responsible for actions in virtue of their relationship to our

²⁰ Raz (n 11), 231.

capacities of rational agency'.²¹ If this were the case, the RFP would avoid the challenges associated with holding agents responsible simply because they acted rationally, able to encompass within the bounds of responsibility actions regardless of whether a particular evaluative position deems them rational or irrational. On closer examination, however, it becomes clear that Raz's principle fails to give meaning to responsibility and the evaluations that follow it.

According to Raz, we are responsible for those omissions that result from a failure of our functioning as rational agents.²² A failure in rational functioning is not necessarily a failure to follow right reason, but a failure of the exercise of rational capacities: forming and executing intentions.²³ It:

is not success consisting in following right reason. It is the successful functioning of our capacities of rational agency. People often do what they should not do without their powers of rational agency malfunctioning. This, for example, can be the case when they have false, but rational, beliefs about what they ought to do.²⁴

A failure in rational functioning is not a failure to have or follow the reasons that are best. It is still a failure to follow reasons, but the reasons that matter are those the responsible person in fact has, rather than those a person ought to have. Raz contrasts failing to set the alarm clock (which I had reason to do this morning) with failing to call a random unknown person in Munich (which I did not have reason to do), noting our

²¹ *ibid*, 227.

²² *ibid*, 248.

²³ *ibid*, 230.

²⁴ *ibid*.

responsibility only for the former.²⁵ On Raz's account, we are responsible for omissions only when the successful functioning of our rational powers would have resulted in us doing what we did not do.

The dividing line for responsibility is not an evaluative question of what reason should be followed, but a descriptive one of what was followed – or not, as the case may be. The RFP functions without requiring evaluative agreement. Raz's demarcation of responsibility avoids evaluative questions. However, the descriptive considerations it applies are, crucially, descriptions of a particular individual's evaluative judgment.²⁶ Raz's responsibility is not defined by reference to shared humanity, but to individual positions.

For Raz, our responsibility depends upon whether our actions are expressions of rationality in following the reasons we have. The result is that responsibility for identical actions can differ between individuals, solely on the basis of their personal values. On Raz's account, responsibility is the product of individual reasons. Responsibility is relative.

The relativism of responsibility prevents it from ever enabling evaluative attitudes across disagreement. If I am talking of my own responsibility, or that of someone whose values I share, then I can work on the basis that I have a reason to set an alarm, and not to call Munich. Consider how the situation changes, though, if it is not my phone call to make, but that of Neville Chamberlain in the mid-1930s. A historian looking back might make the claim that Chamberlain should have made many more calls to Munich than he did, and may point to reasons why he should have done so. Yet Chamberlain did not

²⁵ *ibid*, 248.

²⁶ *ibid*, 230.

himself perceive those reasons, and has accordingly not failed in the functioning of his rational powers. As a result, on Raz's account, he would not be responsible. In hindsight, there were reasons for him to act, but he was ignorant of those reasons. This ignorance absolves him of his responsibility.

Not only Chamberlain's blame, but his responsibility has been absolved by the different reasons adopted then and now. We may not blame Chamberlain, but his potential responsibility presents an altogether different question. When making attributions of responsibility, why should we be limited, as observers, by what the agent we are holding responsible knew? On Raz's account, attributions of responsibility for omissions are rendered relative to the people they describe, and the context they are directed towards.

Beyond omissions, further problems exist. Whether one succeeded or failed in one's rational functioning will depend upon one's reasons, and therefore one's values. This is crucial because it is not the same to be responsible for our failures as it is to be responsible for that which is the product of our rational powers. Only in the latter case can we be deemed to have done something intentionally, for the levels of guidance and control that Raz deems necessary for intention are not compatible with rational failure.²⁷

Raz provides three conditions for control of an action:

'(1) either one performs it because one intends to do so, or one performs it, aware that one does so, by performing another action that one intends to perform; (2) the performance is guided by one's intention and one's beliefs, so that to the extent that one's factual beliefs are true one does not, in performing the action, do

²⁷ *ibid*, 229.

anything else that one believes one should not (on balance) do; (3) in so far as realization of the previous conditions depends on control of one's body they are securely realized.'²⁸

All controlled actions must, on the above account, be either directly intended or the known consequence of an intended action. Not every rational failure is an unintentional action, but no rational success can result in unintentional action. As a result, whilst the boundary between controlled and uncontrolled action is not the difference between having right and wrong reasons, where that boundary lies will undoubtedly depend upon what an agent deemed themselves to have had reason to do. Under the RFP, the divide between intention and negligence is tied to the reasons of the potentially responsible agent. Whether someone is an intentional or accidental actor is a product of the reasons they had. This dependence on individual reasons relativises the boundary between intention and negligence, and has further consequences when taken together with Raz's approach to responsibility for unintentional action.

According to Raz:

'We are non-derivatively responsible for unintentional actions only if they are the results of a failed intentional action that falls within our domain of secure competence. Only then is the action due to a failure of our powers of rational agency, in the meaning of the principle.'²⁹

²⁸ *ibid.*

²⁹ *ibid.*, 244.

In the absence of intention, the only difference between responsible action and irresponsible action is whether that action was the result of a failure of rational agency. To attract responsibility, unintentional action must be the product of what I had reason to do but failed to do. Whether something was a failure of my rational agency is contingent upon my values; the unintentional acts we can be responsible for are once again dependent upon the values we hold. Responsibility is once again relative.

What we have reason to do is a product of our personal and cultural values. In relation to both actions and omissions, responsibility on Raz's model depends upon whether our actions are expressions of rational functioning in following the reasons we have. Responsibility under the RFP can change depending on the values held by a potentially responsible individual, denying the responsibility of individuals not because they are less able to act rationally, but simply because our values differ to theirs. Responsibility for identical actions, or omissions, can differ between individuals.

Responsibility under the RFP is relative: it changes between contexts and between individuals, such that the attitudinal statements it enables convey different things when applied to different people in different contexts. With the divergence of individual reasons comes the relativism of responsibility, and we are prevented from stating what someone is responsible for without reference to who they are and what is important to them.

Rarely will we be able to fully describe the personal and cultural values of any individual. If I attempt to convey attitudes over any cultural or geographical distance, I am left unable to say what a Roman had reason to do in Ancient Rome and therefore

incapable of describing their responsibility.³⁰ Beyond the problem of distance, to the extent that personal reasons are self-created they are often beyond the perception and understanding of others. Attributions of responsibility are altered with differing reasons, whether these changes are the result of distance, or the uniqueness of individual values.

Responsibility becomes a product of individual reasons, and changes for different individuals. There can be no standard of responsibility shared between individuals, as the possibility of attribution collapses into relativism. When responsibility loses its shared applicability, it loses not only its meaning, but also its purpose. If standards of responsibility become dependent upon individual and cultural values then responsibility ceases to enable evaluative attitudes. Responsibility for rational functioning removes the possibility of identifying the common transevaluative ground of our disagreements, by denying the existence of any common interaction with the world. Agents' attributable relations to events, and consequently their responsibility, are rendered relative to reasons.

The only possible escape from the relativism of responsibility is to incorporate the notion of objective rationality into the rational functioning principle. If inherent in rational success is according to an objective standard of rationality, then responsibility accords to this objective standard. By adopting this approach, we can remove the distinction between different kinds of failure noted above. We may also make the distinction between success and failure subject to the proviso that in order to succeed, one must follow objectively right reasons; this objective standard prevents variations across distance.

³⁰ Bernard Williams, 'Relativism, History, and the Existence of Values' in Joseph Raz, *The Practice of Value* (Oxford University Press 2005), 106-107.

Through adding an objective standard of rationality to the criterion for responsibility, responsibility becomes dependent not just upon ‘the ability to realize the normative significance of the normative features of the world’, but also to ‘respond accordingly.’³¹ Agents are required to respond in accordance with their ‘understanding of concepts of true values.’³² An objective standard of rational functioning requires adoption of these true values, to the exclusion of defective ‘false values’.³³

Such an addition requires great confidence in realism, for it leaves responsibility dependent upon the existence of universal moral truths. It moves far beyond a standard realist account by declaring that universal values determine not only what I should do, but also what I can be considered to have done. Most problematically of all, it undermines the function of responsibility by denying the possibility of shared human experience enabling meaningful disagreement. If responsibility is determined by reference to rationality and true values determine rationality, then insofar as we share values, there is no evaluative disagreement for responsibility to enable. Where our values differ, there is no shared attributability that can establish a basis for our disagreement.

In the cases where responsibility is needed – in instances of evaluative disagreement – it cannot function to enable disagreement. Instead, the introduction of true values denies disagreement. If two parties disagree over responsibility, then the party whose evaluative position accords best to true values has made the correct evaluation. There is no disagreement, only right and wrong evaluative positions.

If we are to disagree over objective standards of reason, then we need access to those true values. Evaluative disagreement is possible only when we have perfect moral

³¹ Joseph Raz, *Engaging Reason* (Oxford University Press 1999), 68.

³² Joseph Raz, *The Practice of Value* (Oxford University Press 2005), 24

³³ *ibid*, 24-25.

knowledge. However, if we have such perfect knowledge then, by definition, our evaluations are not in competition, as they are both perfect: reflecting true moral values. There is no evaluative disagreement, and no function for responsibility to fulfill, because attitudes of accountability are already part of our universal consciousness. The extension of realism into the concept of responsibility can only dissolve the disagreements that responsibility enables. Introducing true and false values to a concept of responsibility for rational functioning shields responsibility from relativism, but in doing so denies its purpose.

Responsibility on Raz's account enables the communication of attitudes across evaluative disagreements by basing responsibility not on evaluative questions of what an agent should do, but on the relation between their actions and what they would have done had their rational powers functioned properly. However, the resulting basis of responsibility – personal reasons – is not shared, but a set of personal and often unknowable considerations that vary with cultural context. The relativism of Raz's responsibility renders the attitudes it enables meaningless. Appeal to true values to determine rationality offers no assistance. The impossibility of access to objective standards of reason makes any concept of responsibility defined by reference to them incomprehensible.

Attributing responsibility on the basis of agent characteristics is deeply problematic. The demands of moral luck and reason force both control- and rationality-based conceptions of responsibility into dependence on evaluative judgments and positions. Responsibility attributions lose the shared transevaluative content that allows us to make sense of our evaluative attitudes. A broader conception of responsibility as a

product of rational functioning, whether successful or failed, looks to avoid evaluative dependence. However, adopting an agent's reasons as the determinant of responsibility results in responsibility attributions that are relative to their subject, and equally meaningless, when travelling across evaluative disagreement. There is a need to identify alternative means of defining and demarcating responsibility.

4. Redefining responsibility with a focus on events, not agents

The approaches so far have treated responsibility as something agents attain. In identifying a concept of responsibility that can enable attitudes to travel across evaluative disagreement, it is necessary to look at responsibility from a different perspective. Responsibility need not be a product of something about a person that makes them responsible. It may instead be that something about certain events makes the people involved in them responsible.

A. The responsible agency of intentional human actors

A starting point for redefining responsibility is to look to the paradigm of responsibility: intentional adult human action. Intentional action by adult humans provides a framework of responsibility practices from which we can build a generalised concept of responsibility.

When looking at an event, or events, intentional actors have a special role. Against a backdrop of other causes they are distinguished as actors. Just by being more than causes, actors establish their own intervention in the chain of causation, demarcating

their presence within events.³⁴ They do this simply because they become someone who did something. By acting, they become actors.

Action transforms actors into more than mere causes. Action is elevated above causing, so actors are elevated above causes. In the context of a chain of causation – of causes – the actor becomes the start of something new within that chain.³⁵ Their involvement in the chain does not follow on from previous causes, but represents a break in that chain. With this fresh start the actor in question becomes the origin of the consequence of their acts, whilst a mere cause cannot be responsible for consequences.³⁶

Action, the mere fact of it, says something about the actor: that they are distinct from causes such that we can deem them responsible agents in the context of events. Action is in this sense self-constitutive.³⁷ Importantly, it is self-constitutive in front of an audience.³⁸ Action, then, communicates to an audience. It does not necessarily express anything deeply-held or long-standing about the entity, but it does say that they, as an acting agent, took a special role by trumping causation. It communicates the entity's special status as an agent. A dividing line can be seen between responsibility and non-responsibility: the communication of agency.

The distinction between action and causation is that the former, beyond merely saying what the actors have done, also says something about the actors themselves. It declares their status above the causal and marks them out as a particular sort of entity: as agents. Action declares the presence of an agent within events by holding up that agent as

³⁴ Arendt (n 13), 176-177.

³⁵ *ibid*, 177.

³⁶ Manuel Cruz, *Taking Charge: On Responsibility and Personal Identity* (Richard Jacques tr, Continuum International 2011), 23.

³⁷ Christine M Korsgaard, *Self Constitution: Agency, Identity, and Integrity* (Oxford University Press 2009), 42.

³⁸ *ibid*.

someone, or something, with a particular role. It is this particular role that distinguishes us from dominoes and justifies the imposition of responsibility. Intentional human adult action represents the paradigm case of such responsible agency, but it does not encompass the whole of responsibility.

B. Beyond intentional acts: the limits of agency

In trying to develop a concept of responsibility we can think about other instances in which events are distinguished from mere causal occurrences. We can look to the first-personal recognition of agency that is implicit in Williams' concept of agent-regret. Only some events are such that agents can feel regret for being involved with them. Only some events are apt for Williams' agent-regret.³⁹

Agent-regret is the emotion felt by those who feel a particular connection between their agency and events, such that the event is felt by the agent to be a product of their agency. Williams introduces the example of a lorry driver who cannot react in time to avoid the pedestrian who suddenly appears.⁴⁰ Williams' driver is not in control of his environment, nor is he necessarily rational, but feels regret that his agency is involved in the knocking-down of the pedestrian. The lorry driver's reflection on events will be from a particular perspective, such that he:

³⁹ Williams, *Moral Luck: Philosophical Papers 1973-1980* (n 10), 27.

⁴⁰ *ibid*, 28.

‘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 himself might have prevented it, an agent’s thought.’⁴¹

This particular reflection on events is agent-regret.

Williams’ agent-regret is defined as the regret a person feels towards past events within which he is involved as an agent, regretting them as the products of his own action.⁴² It applies regardless of intention, for we can regret our own accidents just as much as our intentional doings. This, according to Williams, explains why a faultless lorry driver who runs over a child will always have a particular relation to those events, one that he would not have had if he were not the driver.⁴³ Agent-regret connects agents to outcomes not merely as causes, but as participants.

Williams’ truck driver feels regret not simply because he caused the accident, but because of his role in it as the driver. He did not have any control over the consequences of being a driver, or what it as a result meant to be a driver. What he did have was the space within events to have a role in the accident, to be a driver. Before events spiralled out of his control, the truck driver became, within the context of the accident, the driver of the truck. His identity, his role as a participant, was carved into the accident. As a result, ‘there is something special about his relation to this happening, something which cannot merely be eliminated by the consideration that it was not his fault.’⁴⁴

⁴¹ *ibid.*

⁴² *ibid.*, 27.

⁴³ *ibid.*, 28.

⁴⁴ *ibid.*, 28.

The fact of the driver feeling agent-regret is in part determined by his capacity for emotion. It is also dependent upon his values and perspectives; he may not see killing a pedestrian as something regretful. However, the question of whether he feels agent-regret is preceded by the distinct question of whether the accident was something about which he could feel agent-regret. This is a question that does not examine any internal characteristics or capacities of the driver, but instead examines the event as a whole, looking to the driver's role. It asks whether this is an event that is apt for agent-regret. It was only with the particular relationship the driver had to events – that of a 'doer', in this case a 'driver' – that he was able to feel the particular form of regret that is agent-regret. It was only because of the nature of events that regret for being an agent involved in events became possible.

Agents may often feel regret, but only feel agent-regret within a more limited set of circumstances. Williams' driver feels agent-regret because he identifies as the 'doer of the deed'.⁴⁵ Had a different driver been in a cab without a wheel, careering out of control along the same road and hitting the same pedestrian, he may feel deep regret over being in the cab at that time, but this was not regret for being an agent involved in events. Rather, it was regret for being in the wrong place at the wrong time. An event such as this could not give rise to agent-regret, and there is a corresponding absence of any role for the driver in events. There was no space for agency, for becoming an agent was impossible in the absence of a steering wheel.

At the point where the pedestrian stepped out onto the road, it became the case that the pedestrian would die. However, at the point Williams' driver became a driver, at

⁴⁵ Gabriele Taylor, *Pride, Shame, and Guilt: Emotions of self-assessment* (Oxford University Press 1985), 97.

the point he started driving, it was not pre-emptively the case that the pedestrian would die. Thus, whilst there was not space for the driver to change the outcome once he began driving, there was space for agency: for an individual to become the driver who killed the pedestrian.

Only Williams' driver could become a driver who later kills. By contrast, the wheel-less passenger could not be a passenger, or a driver, who later kills. There was no possibility for him to enter into the event as a participant; he never became more than a bystander. He was never in a position to have a role. It is not mere inevitability that is important here, but rather inevitability at the point of involvement of an agent. The role that binds us to events may not initially connect to the eventual consequence of our involvement. Our roles can change, and the removal of control from our hands does not undermine our changing roles: it can only prevent our initial introduction. Once introduced, we may be bound up in events but blameless because of an inability to change things. This is distinct, however, from situations in which we never have the opportunity to become bound up at all.

C. From agency to responsible agents

The relationship between agents and events need not be described by reference to agents' control, or authorship, but instead the roles agents occupy within events.⁴⁶ Responsibility attributions must describe our position in relation to events, but need not be defined by the manner in which we enacted or controlled those events. Agent-regret attaches to those consequences that we helped to bring about, regardless of whether we intended to do so,

⁴⁶ Donald E Polkinghorne, 'Narrative and Self-Concept' (1991) 1 *Journal of Narrative and Life History* 135, 146.

whether the outcome was within our control, or whether we could have done anything different.⁴⁷ Agent-regret is regret for being someone who has done something, for being an agent with a role. It is conditioned upon the position of an agent within events as a ‘doer’, and it is only from this position that agent-regret (as distinguished from simple regret) is possible. This is another instance of the existence and inhabitation of an agent-role, distinguishing events from mere causal occurrences.

Implicit in agent-regret is the having of an ‘agent-role’ in events. Whilst the feeling of agent-regret depends upon our emotional perspectives and capacities, the possibility of agent-regret is a characteristic of events that indicates the existence of a space for agency. Whether agents have this space is a product of the events that occur around them: not of any rationality or control they themselves possess. The space for agency arises out of the way in which we describe events, such that they need a ‘doer’. When events, in the absence of that ‘doer’, would be causally under-determined, that doer inhabits the space for agency. They fill a gap that would otherwise exist within the causal chain, and in doing so take on a particular role as someone, rather than something, within a pattern of events.

That agency is conditioned upon a ‘doer’ role does not imply a requirement of action as a condition of agency. It is quite possible to become someone who did something not by acting, but by failing to act. What matters is not action, but agency, which exists in a range of forms beyond the boundaries of intention and control. The inhabitation of an agent-role takes a range of forms, including omissions.

⁴⁷ Stephen De Wijze, ‘Tragic-Remorse - The Anguish of Dirty Hands’ (2004) 7 *Ethical Theory and Moral Practice* 453, 462.

Agency expressed in terms of a position within events can ground our responsibility for the consequences of that position. Williams' driver's involvement in events was not merely causal.⁴⁸ The possibility of agent-regret demarcates the driver's relationship to what happened, as an agent. His interaction with the world enables our interaction with him, through the communication of evaluative attitudes. The driver may be deemed responsible for the pedestrian's death, though of course this in no way implies his blameworthiness. Responsibility enables attitudes, not simply blame, and a responsibility attribution may lead us to provide a sympathetic, rather than outraged, response.

Agent roles do not only enable agent-regret, but also establish our agency within events. The possibility of agent-regret supervenes particular evaluative stances, existing above and across evaluative disagreement. The connection between agents and events that establishes agent roles is part of, and defined by, our shared human experience. It is constructed from our common perception, grounded in the actuality of human experience, such that the boundaries of responsibility can be demarcated by the structure of our descriptions.

Williams' driver became bound up in events by becoming a driver. Responsibility is a product of his 'doer' status. The boundaries of responsibility are demarcated by the structure of our descriptions. When we move from speaking about causes to agent descriptions, and individuals are said to have driven, or hit, or killed, they are responsible as agents. They are drivers, or hitters, or killers, and responsible on that basis alone. They are to be distinguished from non-responsible individuals who may be no less helpless in controlling outcomes, but whose movements are the consequence of some outside force

⁴⁸ cf R Jay Wallace, *Looking Backward (with Feeling)* (Oxford University Press 2013), 37-38.

that overcomes their agency, inhabiting the space they would otherwise take. When the space for agency is occupied or absent, we lose the ability to be responsible. When we have the space for agency, this implies our responsibility regardless of any rationality or control we may possess.

The agent-role of agent-regret demonstrates agents' capacities to make, be the subject of, and understand responsibility attributions – without requiring that responsible agents possess particular characteristics in order to do so. The space for agency arises simply because the description of events would be insufficient without the agent's presence. Without agency, causation would be under-determined; there would be something missing. When agents occupy this space, they can be held responsible on the basis of the descriptions of the pattern of events around them.

The possibility of agent-regret that accompanies such instances of responsible agency emphasises and enables the communicative role of responsibility. With the possibility of agent-regret comes the possibility of an agent understanding his or her own position within events. The first-personal self-reflection of agent-regret reveals the possibility of understanding attributions of responsible agency communicated by others. The capacity for understanding one's own responsible agency is a consequence of the agency proscribed by the possibility of agent-regret, and it enables the shared understanding of attributions of responsibility. An agent with a role apt for agent-regret is not just as someone who has done something, but someone who can be the subject of responsibility attributions.

Just as the space for agency permits us to engage in self-reflective agent-regret, so too does it allow us to express evaluative attitudes towards others, and to understand

those attitudes when they are directed towards us. Responsibility framed as a counterpart to agent-regret captures the differing possibilities for adults, infants and dominoes to have roles and attract responsibility. It does so whilst fulfilling the function of responsibility, invoking shared transevaluative content and in doing so endowing us with the tools necessary to convey accountability across evaluative boundaries.

Responsibility attributed on the basis of the possibility of agent-regret provides a shared, mutually understandable reference point for disagreement. This is responsibility that takes its justification from the particular role of the agent within events. Like events that result from intentional action, those that are apt for agent-regret demonstrate the special significance of the agent as something more than causal: as someone responsible.

Such an approach recognises the evaluative significance of responsible agency, without resorting to evaluative dependence. The possibility of agent-regret is a product of the description of events, described in terms of shared basic concepts like driver, hitter, and killer. Agent-regret is a product of our shared human experience, arising out of common definitions and understanding. The space for agency exists because there is causal space for an agent to have a role. This role is constructed from transevaluative content, such that Williams' truck driver's plight may be understood from any evaluative perspective. We can attribute responsibility so long as we can convey to an agent their role as an agent; the limits of responsibility coincide with the possibility of communication of agency.

Linking responsible agency to the rationality or control of the agents in question offers no satisfactory means of delineating the boundaries of responsible agency. The solution is to look not to the characteristics possessed by the agent, but instead to the

character of the event they are involved in. It is to ask: was there space for agency? Was there space for someone to take a special status, or was this simply a case of x causing y causing z ? Responsible agency exists where there is space, within the narrative of events, for such agency. Where agent-regret is possible, responsibility may be both communicated and understood. Responsibility can provide the shared transevaluative content needed to communicate evaluative attitudes.

Responsibility does not in this way eliminate disagreement, or detach itself from evaluation. Responsibility attributions enable evaluative disagreement. They are a fundamental component of our evaluations, but crucially are transevaluative prerequisites of our attitudinal statements. They ensure that attitudes of accountability are understandable across boundaries of disagreement. Responsibility arising from the possibility of agent-regret offers a means of recognising the shared content of evaluative disagreements.

Responsible agency is demarcated by the possibility of agent-regret. That an event is apt for agent-regret demonstrates the inhabitation of space for agency, such that we can become someone who did something, and the shared communicability of attributions of responsibility, such that we can become someone responsible for doing something. The possibility of self-reflective agent-regret enables the understanding of responsible agency. Responsibility attributions can be made, and accountability communicated, by reference to shared human experience.

5. Responsibility reconsidered by reference to purpose

The attitudes of accountability are fundamental to the relationships between agents. Responsibility is a means of grounding our attitudinal utterances in universally understandable attributions of agents' relationships to events. When assessed against this aim, attempts to define responsibility by reference to capacities of rationality fail to capture responsibility without reliance on evaluative stances, defeating the purpose of responsibility and leaving the attitudes we express either incomprehensible or meaningless. Basing responsibility on control forces us to make judgments on how much, and what kind, of moral luck can be accepted. In doing so, it renders responsibility evaluatively dependent, undermining its function as a basis for our evaluations. A third approach is to look not to the capacities of agents, but to the roles they play. These roles carry inherent significance in demarcating agents as agents, distinguished from causes by having something said about them beyond simply what they have done.

Agent roles justifying responsibility coincide with the possibility of agent-regret. They exist only within the space for agency, a characteristic of events – not agents – that allows for the introduction of an agent as more than a cause. Within this space we find responsibility unbridled by agent characteristics and disconnected from evaluative positions. With this meaning, we find responsibility with purpose.