

## WHAT IS A RESPONSE TO WRONGDOING?

Nicolas Nayfeld<sup>1</sup>

<sup>1</sup> Jean Moulin Lyon 3 University, France

Original scientific paper – Received: 10/03/2025 Accepted: 11/08/2025

### ABSTRACT

This article starts from the assumption that to properly assess the merits of “response retributivism”, we must first clarify the nature of a response to wrongdoing and how it differs from a mere reaction. I propose a communicative account of responses, arguing that a response to wrongdoing is a distinctive form of action motivated by the wrong, whose special feature is that it is addressed to the wrongdoer, has a confrontational-communicative dimension, and needs to be identified by the wrongdoer as a response to his wrong. I argue that this definition allows us to rethink the concept of responsibility, to make progress on the debate whether there is a duty to respond or react to wrongdoing, and to refocus the discussion toward what Strawson calls “reactive practices” as opposed to “reactive attitudes”.

**Keywords:** response retributivism; wrongdoing; communicative account; responsibility; reactive practices.

## Introduction

In her 2020 article, Leora Dahan Katz defends a form of retributivism called “response retributivism” and argues that

[w]hile moral norms generate primary duties to act and refrain from acting in particular ways (primary duties), (...) such norms also generate secondary duties to *react* and *respond* to violations of primary duties in particular ways. (Dahan Katz 2020, 2)

This theory is refreshing because it retains the deontological core of retributivism while expanding the focus beyond punishment, suggesting that punishment is only one possible response to wrongdoing among others.

However, the problem with this theory is that it relies on a fundamental ambiguity as the quote above shows: it remains unclear whether Katz’s account implies a duty to *respond* to wrongdoing or a duty to *react* to wrongdoing, or both. This initial problem is compounded by a second: the distinction between responding and reacting, as such, has been neglected by philosophers. They generally do not define the concept of response, taking it for granted and often equating it with Strawsonian “reactive attitudes”.<sup>1</sup> Mainstream psychology, for its part, has seized on this distinction, but the way it draws it seems implausible to me: reactions, it is said, are impulsive and unthinking, while responses are thoughtful and deliberate.

This paper has three related aims. The first and main objective is to clarify this distinction and to offer what can be called a communicative account of responses insofar as it shares much with communicative theories of punishment (particularly Duff’s) and censure theories (especially Narayan’s)—though it extends their insights to responses more broadly, beyond punishment and censure. I should warn from the start that I will not try to determine the necessary and sufficient conditions for a response/reaction to occur. Rather, I will attempt to point out the most salient features.

The second objective is to show that such clarification allows us to make significant progress on the debate whether there is a duty to react or respond to wrongdoing.

---

<sup>1</sup> For example, although her book is entitled *Forgiveness and Retribution: Responding to Wrongdoing*, Holmgren (2012) does not, to my knowledge, say what it means to respond to wrongdoing.

The third objective is to show that such clarification can help understand what Strawson means by “reactive practices”, a surprisingly understudied notion (see Metz 2008, for an exception), in contrast to “reactive attitudes” on which much ink has already been spilled.

I proceed in six steps. First, I introduce three cases (and variations of each) to make the distinction between not reacting, reacting, and responding intuitive. These cases (“the battered woman”, “sexual harassment on the subway”, “the unruly young child”) serve as a basis for the entire analysis. Second, I identify commonalities between reactions and responses, emphasizing that both involve behavioral action and stand in opposition to inaction, which means that neither letting go nor mere emotions count as reactions or responses within this framework. Third, I explore the differences between reactions and responses, focusing primarily on the specificity of responses, the crucial difference being that responses are *addressed* to the wrongdoer. I explain why this notion of “addressing” is complex and requires careful unpacking. Fourth, I examine the relationship between responses and responsibility on the one hand, and the specificity of penal responses on the other. Fifth, I address a potential objection: that my analysis might overlap with Strawson’s distinction between the objective attitude and the participant attitude or between objective attitudes and reactive attitudes. I argue that this objection is misplaced; rather, a more apt comparison can be drawn between what I call “responses” and what Strawson calls “(reactive) practices” in the opening and closing sections of “Freedom and Resentment”. Finally, I explore the normative implications of this conceptual framework and argue against the idea of a duty to respond to wrongdoing.

## 1. Three cases

I will start with three cases (and their variations) that vividly illustrate the distinction I want to draw between reacting to a wrong and responding to it. I do not think that the distinction I am going to propose between response and reaction corresponds perfectly to ordinary usage. The distinction between response and reaction in ordinary language is not so clear-cut, and the two terms are sometimes used interchangeably. However, I believe that the distinction I want to draw is *fairly intuitive* in French (my native language) and in English (as native speakers have confirmed to me), that it *corresponds partially to ordinary usage*, and therefore *is not entirely stipulative*. To be clear, in this article, I am not interested in ordinary language as such, but rather in the contrast between different types of situation or action (labelled A series, B series and C

series below) that exhibit distinctive features.<sup>2</sup> In fact, it often happens that valuable philosophical distinctions overlap only partially with ordinary language (see for instance Kant's distinction between belief, knowledge and opinion in his *Critique of Pure Reason*).<sup>3</sup>

#### Case 1: the battered woman

1.A. A woman who has faced domestic abuse for years remains silent, refraining from any action out of fear of her husband. She bears her suffering, hoping her ordeal will eventually come to an end.

1.B. Another woman in a similar situation contacts a domestic violence hotline. She works with volunteers to plan her escape, carefully concealing any evidence of her call to protect herself from her husband's potential retaliation.

1.C. Another woman, after years of abuse and a final assault, uses a shotgun to kill her husband, then calls the police to confess (a scenario reminiscent of the Jacqueline Sauvage case in 2012).

In 1.A, the abused woman experiences reactive feelings/sentiments/emotions, but in a sense yet to be determined, she neither reacts nor responds to her husband's abuse. In 1.B, the abused woman reacts to her husband's abuse by making a safety plan, though she does not respond directly to it. In 1.C, the abused woman responds violently and irrevocably to the assault she has just suffered, but also to the years of abuse she has endured.

#### Case 2: sexual harassment on the subway

2.A. A bystander witnesses sexual harassment in a subway station but does nothing, held back by fear and discomfort.

2.B. Another bystander in the same situation discreetly reports the incident to law enforcement, leading to a swift police intervention.

2.C. A third bystander, upon witnessing the harassment, confronts the perpetrator directly, loudly condemns the behavior, and warns of further consequences if the harassment does not stop.

In 2.A, the bystander fails to react, although the situation does not leave him emotionally indifferent. In 2.B, the bystander reacts by reporting the event through an established channel. In 2.C, the bystander himself provides a response to the harm he witnesses.

---

<sup>2</sup> For a similar methodological insight, see Hart (2012, v).

<sup>3</sup> I would like to thank reviewer 1 for prompting me to clarify this point.

### Case 3: the unruly young child

3.A. An 18-month-old child pulls a book from the family bookshelf and starts tearing it. The parents decide to ignore the incident, thinking “it’s no big deal”, “this book is bad anyway”.

3.B. Different parents in the same situation choose to childproof the bookshelf by moving the books out of reach.

3.C. Other parents scold the child, explaining that tearing books is unacceptable, attempting to instill a sense of respect for property.

In 3.A, the parents opt for inaction, choosing not to react. In 3.B, the parents react to prevent future incidents but avoid involving the child directly. In 3.C, the parents feel it is appropriate, even necessary, to respond to the incident, and therefore provide (what they believe to be) an educational response to the incident.

Now that we have illustrated the distinction, we can begin to analyze it. I will first explore the commonalities between B series (which I classify as reactions to a wrong) and C series (which I classify I responses to a wrong).

## 2. Commonalities

Both reactions and responses are motivated and justified by a wrong (an offense, a loss, a harm, a slight, etc.).<sup>4</sup> In B series and C series, the answer to the “why?” question (Anscombe 2000) will mention the wrong. Why did you call the police? *Because he was harassing her*. Why did you start yelling at him? *Because he was harassing her*. A child can push his brother into a puddle because he said something mean to him (response), but also push him “for fun” (not a response) (see Metz 2008, 228).

Both reactions and responses can vary along a spectrum from impulsive and disproportionate to thoughtful and measured. The common assertion, particularly in psychological and self-help literature, that reactions lack any thought while responses are thoughtful is, I argue, mistaken.<sup>5</sup> In anger, a person might respond to an insult with a slap. Conversely, a person might

---

<sup>4</sup> As my brackets indicate, I use the term “wrong” as a generic term: I am interested in actions that should not have been done and for which someone is accused. Although we can also respond to a benefit, the distinction between responding and reacting does not seem to work as well with benefits.

<sup>5</sup> See <https://www.psychologytoday.com/us/blog/the-power-prime/202110/the-difference-between-reacting-and-responding>

react with great composure and professionalism to threats from an inpatient in a psychiatric hospital by calmly restraining him.

Both reactions and responses can take place under social pressure. Public figures such as politicians and judges are expected, even required, to react and respond to wrongdoing as quickly as possible. Failure to do so can attract public criticism. This social pressure has arguably intensified with the rise of social media. In fact, Dahan Katz's argument (mentioned in the introduction) in favor of a duty to react and respond illustrates this growing expectation.

Both reactions and responses need not originate with the victims of wrongdoing (if any). They may instead come from third parties, such as bystanders (as in cases 2.B and 2.C) or the state (for instance, when it intervenes to punish criminal wrongdoing). The same is true of linguistic responses (answers): Mary's parents, for example, might respond to Paul's letter by instructing him to cease contact.

Finally, both reactions and responses in B series and C series are behaviors or clusters of behaviors. In case 2.A, although the bystander may experience guilt, he fails to react: as the word suggests, to re-act, you have to act. This is why forgiveness only sometimes qualifies as a response; private forgiveness (e.g., overcoming resentment without any interaction with the offender) is not a response, whereas overt forgiveness (e.g., telling the offender "I forgive you") is. Some may argue that "not resisting" in a situation (such as a bank employee complying calmly with robbers) constitutes a reaction. But this is because "not resisting" in this case involves active behaviors: remaining composed, cooperating, and prioritizing safety, rather than taking no action at all. In fact, both reactions and responses are opposed to inaction in all its forms, whether reflective or unreflective. Deciding to "let go" in the face of wrongdoing, for instance, does not constitute a response; rather, it involves a choice to abstain from responding. Brunning and Milam (2022, 6) give examples of such a choice: a hate-speech victim feels further blame would be pointless and continues on his way; a friend tolerates another's rude behavior to avoid continuous confrontation. Letting go, then, is an option we may choose—or even feel compelled to choose—in the face of wrongdoing, but it does not qualify as a response in the strict sense.

### 3. Differences

Now let us look at the differences. Responses, as illustrated by the C series above, are always *addressed* to someone, always have an *addressee*—

namely the wrongdoer. If an individual commits mass murder in a school but then commits suicide, there is no way to respond to his crime. However, the authorities can react by imposing stricter gun laws, increasing security in schools, improving emergency evacuation procedures, etc. Reactions need not be addressed to the wrongdoer, though they can address the causes of wrongdoing. Another example: when the identity of the wrongdoer is unknown, it is impossible to respond to his wrongful conduct, for instance by censuring him, “since censure is condemnation addressed to the perpetrator” (Narayan 1993). However, it is perfectly possible to react to his wrongful conduct, for instance by expressing disapproval of it, by denouncing and disavowing it (Narayan 1993).<sup>6</sup>

This means that responses inherently possess a dialogical, face-to-face, “confrontational”<sup>7</sup> (Shoemaker 2015, 112) quality, which is absent in the B series above (reactions). When responding to a wrong, I engage the wrongdoer in a conversation.<sup>8</sup> Or perhaps one should say that I *continue* the conversation initiated by the wrong.<sup>9</sup> As Murphy puts it (followed on this point by Hampton<sup>10</sup> and Hieronymi<sup>11</sup>):

[I]njuries are also messages—symbolic communications. They are ways a wrongdoer has of saying to us, “I count but you do not,” “I can use you for my purposes,” or “I am here up high and you are there down below.” Intentional wrongdoing insults us (...). (Murphy 1988, 25)

Responses can be a one-way process or a two-way process (Duff 2025, 93-94): it is a one-way process when no response is expected from the wrongdoer or when the intention is to terminate the conversation immediately (as in case 1.C when the victim kills the wrongdoer), and it is a two-way process when the response seeks a response (as in cases where we expect, or explicitly demand, an apology). By contrast, reactions can

---

<sup>6</sup> Narayan’s interesting distinction between censure and denunciation (expressive theories) can be seen as a particular instantiation of the more general distinction I propose between responses and reactions. The distinctive characteristics she gives overlap largely with my own.

<sup>7</sup> It is also the word employed by Narayan in her justification for censure, by opposition with mere denunciation: “We are interested in *confronting* the perpetrator with the judgement that she is a responsible wrongdoer (...)” (Narayan 1992, 172).

<sup>8</sup> In general, my analysis has much in common with McKenna’s “conversational theory of moral responsibility” (McKenna 2012).

<sup>9</sup> “Responding to an agent’s action (...) is like engaging in an unfolding conversation with the agent whose act can be thought of as the initiation of a conversation” (McKenna 2012, 213).

<sup>10</sup> “By victimizing me, the wrongdoer has declared himself elevated with respect to me, acting as a superior who is permitted to use me for his purposes. A false moral claim has been made. Moral reality has been denied” (Hampton 1988, 125).

<sup>11</sup> “[A] past wrong against you (...) makes a claim. It says, in effect, that you can be treated in this way, and that such treatment is acceptable” (Hieronymi 2001, 546).

leave the wrongdoer completely out of the conversation (“as if her participation in any discourse about the wrongdoing (...) is pointless” says Narayan 1993, 172). In fact, reactions can be seen as a way of avoiding or bypassing a conversation that we do not want to have for various reasons, including “the strains of involvement” (Strawson 2008, 10).

It is worth noting that the confrontational aspect of responses does not mean that they are always retaliatory. The concept of response should not be confused with the concept of retaliation/reprisal. The noun ‘retaliation’ comes from the Latin verb *retaliare* (to pay back in kind), which is also found in the *lex talionis* (an eye for an eye, a tooth for a tooth): retaliation thus implies a form of paying back: I hurt you because you hurt me. In contrast, responses to a wrong can be gentle and pacifying: we can respond to a wrong with dialogue, without intending to harm the wrongdoer.

The confrontational nature of responses also explains why we may be reluctant to respond to wrongdoing in certain circumstances: we may not want to respond, for instance, to racial harassment either because we do not see the point of confronting the wrongdoer (who, we are sure, will never understand why what he did was wrong), or because we think confronting him would be dangerous (risk of escalation).

At this point, at least three apparent counter-examples might be put forward: conviction in absentia, transitional justice, and blaming/forgiving the dead. Let us take them in turn.

In 1940, in occupied France, the Pétain regime sentenced de Gaulle, then in exile in England, to death in absentia for treason. Does this show that (penal) responses need not be confrontational? Not really, because in the case of a conviction in absentia, the conviction is addressed to the culprit (and even beyond to the public at large, as I explain below): it is made public so that he can learn about it (even if his whereabouts are sometimes unknown). The Pétain regime knew that de Gaulle was in England and would be informed of this death sentence. It acted more or less like a person searching in the dark for someone hiding and threatening “If I find you, you will regret it”.

Transitional justice also appears to be a counter-example. The Algerians, for example, have long demanded an apology from France for 132 years of colonization, for the massacres in Setif on May 8, 1945 that left thousands dead, for the use of torture during the Algerian war, etc. Demanding an apology seems to be a form of (non-violent) response to wrongdoing. But the problem is that many of the key perpetrators are deceased. To this we can answer that the Algerians are demanding an



apology from the French state, which has not ceased to exist since the commission of these crimes and whose continuity has been ensured by the succession of different governments. In other words, the addressee of this demand is the French state or its representatives. It is up to them to acknowledge the wrongdoing that gave rise to this demand. On the other hand, if France were to be wiped off the map as a result of an invasion, Algeria would not be in a position to respond to the crimes committed by France and, above all, to demand an apology.

As for blaming/forgiving the dead, this can be seen as a response to the extent that we believe that they somehow hear us (the rationality of this belief is not our concern here). We could also simply say that “[t]hese are derivative cases to be accounted for in terms of how we would respond to the dead were they still alive” (McKenna 2012, 177).

Responses not only have an addressee, but they also convey a message. In other words, they are “communicative” (Macnamara 2015). When I merely react to a wrong, as the B series shows, I do not try to tell the wrongdoer something, to make him understand something. I mainly try to stop the wrong, to prevent it from happening again. What is said by a response—the message conveyed—depends, of course, on the nature of the response, although they probably all express at least some judgment of disapproval, such as “What you did was (what you are doing is) wrong”. For instance, if you respond to a wrong (e.g. infidelity) with forgiveness, telling the wrongdoer “I forgive you”, you probably mean that *although* what he did was wrong, you do not hold it against him.

Because responses are communicative, they can be misinterpreted or difficult to interpret. Consider the example of Jesus turning the other cheek in the Bible. This is a very baffling and unexpected response to a slap. What is it supposed to mean? Perhaps something like “I am against violence: you can slap me as much as you want, I will not respond with violence, I will not play your game”, but other interpretations are possible, as theological controversies show.

Because responses involve an addressee, a dialogical dimension, and a communicative function, they are generally not directed at entities that lack understanding or *logos*, such as trees or flies (Macnamara 2015, 2, 17). Responses “make sense only on the assumption that the other can comprehend the message” (Watson 2019, 230). The more limited the other’s level of understanding, the more limited the possibilities of response. That is the reason why we cannot respond to the aggressive behavior of a young child as we can to that of an adult.

By contrast, we can react to damage caused by entities that lack understanding. Although you will not respond to rat damage in your kitchen (by getting even with the rats or forgiving them), you may react by calling a pest control company; although you will not respond to rain damage in your ceiling (by blaming the rain), you may react by cleaning out your gutters.

We can also react to harm caused by people whose understanding is severely impaired. In the recent Sarah Halimi case in France, the man who killed this elderly woman was declared not guilty by reason of insanity and was consequently sent to a psychiatric hospital and subjected to “security measures” (*mesures de sûreté*) for the next twenty years. The state, or more precisely the criminal justice system, reacted to this crime (took therapeutic and security measures) but did not respond. That is why many people were dissatisfied with this decision: they wanted Kobili Traoré to be sentenced to life imprisonment; they wanted a strong and severe response, and they only got a proportionate and rational reaction.

One might object that, just as we sometimes get angry at an inanimate object, say a chair, after bumping into it, so too we sometimes respond to a “wrong” caused by nonhuman beings. Seneca gives the example of Cyrus the Great who tried to ford an Iranian-Iraqi river, but whose chariot was swept away by the current. Outraged, he had 180 channels dug, dividing the river into 360 streams (Seneca 2012, 80-81). However, in this example, Cyrus the Great punished the river as if it were a responsible agent that had shown insolence. He taught it a lesson as if it could understand the lesson:

We tend to think that we have a right to expect ‘respect’ and cooperation from the inanimate objects that serve our ends, and in the moment we react as if they were bad people, since they clearly are not doing ‘their job’ for us. (Nussbaum 2016, 18-19)

Such conduct is, of course, childish, not to say pathetic, as Seneca points out. It shows how anger can lead to completely irrational responses.

Responses and reactions do not have the same success conditions.<sup>12</sup> What I mean here by “success” needs to be clarified. An act of revenge can succeed at being a revenge, at avenging for instance my dead brother (logical/conceptual success), but it can also succeed at accomplishing various aims: deter, make suffer, and so on (teleological/instrumental

---

<sup>12</sup> I would like to thank reviewer 2, who greatly helped me clarify this entire section.

success).<sup>13</sup> I am not here concerned with the conditions for teleological success, but only with the conditions for logical success: why some of our attempts at responding to a wrong fail to get beyond the attempt stage? The thesis I will defend is that for our response to succeed in the logical sense, it has to be identified as such, to be received as a response by the wrongdoer. The wrongdoer must *understand* that our response is a response *to the wrong we attribute to him*—he does not have to believe that he has committed a wrong, as long as he is aware that we think he has (this means that moral disagreement is compatible with responses). For instance, if someone takes revenge for a wrong by attacking the aggressor, but the aggressor has lost his memory due to a degenerative disease and does not remember the wrong he committed, the revenge cannot succeed, the response fails. Here is another example: if I demand an apology from a person, but that person asks me why, my response—the demand for an apology—has failed. In contrast, reactions to wrongdoing do not have to be identified as such by the wrongdoer in order to succeed as reactions; they operate independently of the wrongdoer’s awareness and understanding.

A counterexample here might be drawn from the film *Old Boy*, where the main character, Dae-su, is kidnapped and wakes up in a sealed hotel room without knowing why. He is released 15 years later, but still ignores why he was kidnapped. He later understands that this torture was part of Woo-jin’s revenge for some revelations he made in high school, which had disastrous consequences. But let us assume for the sake of the argument that he never learns why he was kidnapped. In that case, can we say that Woo-jin responded to Dae-su’s wrong? From Dae-su’s point of view, this would not be a response; or maybe he would speculate that it is a response for something he did, though he does not know what, or that he is the victim of a strange experiment without his consent. But from Woo-jin’s point of view? My intuition tells me that his revenge would not be fully satisfying in this scenario: if Dae-su believes that his torture is not a response to his revelations, but just an experiment planned by the government or a secret society of psychologists, i.e. if he has *false beliefs* about the reasons why he is tortured, the revenge somehow fails in the logical sense—it is an attempted but unsuccessful revenge. This is probably why, in any revenge scenario, the avenger makes sure, or takes it for granted, that the offending party understands that he is getting even.

---

<sup>13</sup> Duff seems to rely on this distinction, though he does not make it explicit: “Communication seeks a particular kind of response from its addressee: if you do not hear or understand me, my communication has failed. It often also seeks not just understanding, but acceptance: I intend, or hope, that you will accept what I communicate, and my enterprise has been to that extent a failure if you are unpersuaded; but so long as you understand, I communicate successfully” (Duff 2025, 94). In my terminology, the understanding condition is a logical success condition: my attempted communication fails at being a communication if you do not understand me. However, the persuasion condition is a teleological success condition: communication did not achieve its objective if you are unpersuaded.

It is worth noting that Aristotle makes a very similar observation in his *Rhetoric*. He explains that we do not punish people when we know that they “will not perceive who is the cause of their suffering and that it is retribution for what [we] have suffered”. He then uses the example of Odysseus and Polyphemus the Cyclops. After gouging out Polyphemus’ eye and rejoining his ship, Odysseus reveals his true identity. This seems to be an act of hubris on the part of Odysseus, but Aristotle thinks that it is perfectly understandable “since [Odysseus] would not have been avenged if [Polyphemus the Cyclops] had not realized both from whom and why revenge came” (Aristotle 2007, 123, 1380b).

Cases in which the response is not identified as such should be distinguished from cases in which the response is identified as such but is completely ignored. For instance, the young child in 3.C may respond to the parents’ scolding by smiling or even laughing, pick up a new book and start to tear it up. In 2.C the harasser may ignore the bystander’s threats, act as if the bystander does not exist, and continue to harass his victim. One last example is what Duff calls the “defiant offender”: “the offender that will not even listen to the moral message that his punishment seeks to communicate” (Duff 2001, 123). In all these cases, the response did not have the effect we wanted it to have (which was to stop the wrongdoing, or to engage the wrongdoer in a serious consideration of what he has done), even though a response proper took place—teleological failure, but logical success.

As we can see, responses to succeed in the logical and teleological sense depend largely on the wrongdoer. Which is not surprising: a conversation can succeed in both senses only if each speaker is able to discuss and is open to discussion. Which also means that responses are vulnerable and need to be backed up by reactions: in 2.C, faced with this failure, the only thing left to do is call the police; in 3.C, faced with this failure, the only thing left to do is move the books out of reach.

My account of responses, though is incorporates insights from expressivism (most notably, Feinberg’s 1968 expressive account of punishment), should be distinguished from the latter; the fact that your reaction to a wrong expresses, say, your hatred does not suffice for it to be a response in the strict sense, since the expression of hatred may be devoid of any confrontational-communicative dimension and lack the identification component described above. As Duff—who also wishes to distinguish his communicative theory of punishment from expressive ones and to whom I owe this point<sup>14</sup>—puts it:

---

<sup>14</sup> For a similar point, see also Narayan (1993).

I can express, and gratify, my hatred by harming the person I hate, and it might not matter to me whether he knows that I did this: I sabotage my hated enemy's car, and am gratified when he crashes, even if he believes the crash was accidental. Sometimes, however, it matters that the object of my expressive action knows about it: I want him to realize not simply that he is harmed, but that I have harmed him. (Duff 2025, 93)

Now that I have highlighted the main differences between responses and reactions, we can more confidently assess what qualifies as a response and what does not. For instance, in her list of responses other than punishment, Walker includes commemoration (Walker 2012, 10), yet, I would argue that commemoration fits more accurately within the category of reaction, that is within the B series. Commemoration bears all the distinctive marks of a reaction: it is opposed to indifference or inaction; it aims to prevent the repetition of the past; it is seen as a civic duty (like calling the police upon witnessing sexual harassment, as in case 2.B); we can commemorate natural disasters (Japan, for example, commemorates the Fukushima disaster) when we cannot respond to them (unless we personify nature as responsible and guilty); it does not need to be recognized as such by the perpetrators of the tragedy being commemorated; it does not imply a communicative confrontation with the author, but rather a communion between us: it is addressed above all to the younger generations.

Although reactions and responses are different, the boundary between them is not always clear-cut. In fact, the same behavior can be considered a reaction or a response depending on factors like context or the agent's mental state. For instance, cutting ties with someone can be understood in two different ways. As a response, it might be intended as a form of punishment, aiming to impose distress or teach a lesson, either as an end in itself or as a means to achieve some other outcome (it thus fits within the C series). In contrast, cutting ties might also be a reaction—an act of self-protection in the face of a toxic relationship, where the painful effects on the other person, though anticipated, are not the intended goal but rather an unfortunate byproduct (it thus fits within the B series). In this second case, the separation tends to be more abrupt and definitive. When cutting ties is intended as a punishment, there is at least some underlying communicative intent, which, paradoxically, maintains the ties by cutting them. This nuanced difference may help explain why so-called *pervers narcissiques* (a French term for manipulative narcissists) often seem to prefer being punished rather than shunned.

## 4. Applications

### 4.1 Responses and responsibility

Some authors have pointed out that the word “responsibility” originates from the Latin *respondere* which, before meaning to answer questions, meant to answer accusations. Thus, etymologically, responsibility means “answerability”:

[A] person who is responsible for something may be required to answer questions (...). To say that a minister is responsible for the conduct of his department implies that he is obliged to answer questions about it if things go wrong (...). The original meaning of the word “answer” (...) was not that of answering questions, but that of answering or rebutting accusations or charges, which, if established, carried liability to punishment or blame or other adverse treatment (...). (Hart 2008, 265)

But responsibility can also be associated with “responses”, not in the sense of verbal answers to questions/accusations, but in the sense of actions motivated by a wrong and addressed to the wrongdoer. In other words, it can also be defined as “responseworthiness”.<sup>15</sup> According to this definition, the word “responsible” parallels the word “desirable” (worthy of being desired). To be responsible for a wrong is “to be eligible to a range of (...) responses with a built-in confrontational element” (Shoemaker 2015, 87).

Thus, an important difference between responsibility as “answerability” and responsibility as “responseworthiness” is that in the first case verbal answers go from the agent to others (the person who is responsible is required to answer questions, accusations), while in the second case behavioral responses go from others to the agent (the person who is responsible is liable to responses such as blame, “moral protest”, punishment, forgiveness, etc.).

Although distinct, answerability and responseworthiness, at least in the criminal justice context,<sup>16</sup> are related as follows: “a person who fails to rebut a charge is liable to punishment or blame” or some other response

<sup>15</sup> “[I]n the broad sense, moral responsibility for an action is a matter of *deserving a moral response* on the basis of the action” (Copp 1997, 452).

<sup>16</sup> I say “at least in the criminal justice context”, because outside of that context we may respond to what we perceive as a wrong without giving the addressee an opportunity to defend himself. When McKenna says that punishment is “a response that might be regarded as fitting *after* a morally responsible agent has had an opportunity to offer an account of her conduct” (2012, 91), he is making a normative claim about when punishment should be inflicted.

“for what he has done, and a person who is liable to punishment or blame” or some other response “has had a charge to rebut and failed to rebut it” (Hart 2008, 265). In other words, the answerable person becomes a responseworthy person if, when accused of a wrong, he is unable to defend himself with a compelling justification or excuse. And if he could not defend himself with a compelling justification or excuse, but acknowledged the wrong done and apologized, he is worthy of a merciful response.<sup>17</sup>

One final remark: the fact that a person is responsible for a wrong in the sense of responseworthiness does not mean that we should respond to his wrong, nor that failing to respond and merely reacting or resenting him would necessarily be blameworthy or inappropriate. It rather means, to borrow Shoemaker’s term cited earlier, that the person is eligible to responses—that they are grounds for punishing, blaming, demand an apology, etc. But of course, people do not always get what they are eligible to—grounds for responding can be defeated or overridden by other, competing considerations.

## 4.2 Penal responses

Penal responses constitute a subset of responses and this raises the question: what is specific about them? It might seem that penal responses, as their name suggests, are defined by their punitive nature, implying a narrower scope than responses in general. However, this assumption is incorrect. The criminal justice system sometimes employs non-punitive responses, such as issuing a legal reminder, granting an absolute discharge (*dispense de peine* in French law), or mandating participation in rehabilitative programs, such as those for substance abuse in drug courts.

In fact, penal responses are characterized not by a single feature, but by several. First, they are structured rather than diffuse; they are administered by an organ of the state—namely, the judiciary (Durkheim 2013, 55). As a result, penal responses involve “coercion that is unilateral only”: the guilty “is subjected to an external duty to which he, for his own part, may offer no resistance”, and this may offend his “feeling of honor” or suspend “his dignity as a citizen” (Kant 1991, 168, note). A convicted person, of course, may appeal, but if the conviction is upheld, he must submit to it, he cannot resist it. Besides, the target of penal responses (what they respond to) is narrower than the target of responses in general: they respond only to

---

<sup>17</sup> The question of the conditions for responseworthiness, like the question of the conditions for blameworthiness, is a complex one that I cannot resolve in this paper. However, it seems plausible that there are both general conditions for being responseworthy and more specific conditions for being worthy of this or that type of response.

criminal wrongdoing (offenses), not to wrongdoing in general. Finally, the addressee of penal responses is not only the offender, but also the public at large.<sup>18</sup> Certainly, some cases of private revenge involve a kind of extended communication: revenge is addressed not only to the aggressor, but also to those around him who are warned “Do not imitate him”. It can also be found in classrooms where the teacher’s responses to disciplinary offenses are addressed not only to the offender, but also to the entire class. However, penal responses go further: for instance, when Harvey Weinstein was sentenced to 23 years, society as a whole was reminded not only of the offense’s gravity but also of the legal and moral condemnation of sexual violence.

Interestingly, some proponents of retributivism or “just deserts” paradoxically defend one penal response, legal punishment, as more humane than therapy or reform, which they view as degrading, i.e. bestializing, infantilizing, or demonizing (Waldron 2010, 282-83). Consider, for instance, the following excerpt from “The Humanitarian Theory of Punishment” by Lewis:

To be “cured” against one’s will and cured of states which we may not regard as disease is to be put on a level with those who have not yet reached the age of reason or those who never will; to be classed with infants, imbeciles, and domestic animals. But to be punished, however severely, because we have deserved it, because we “ought to have known better”, is to be treated as a human person made in God’s image. (Lewis 1987, 151)

Lewis holds that punishment treats the punished person “as a human person” probably because it is a response: no matter how serious the crime, punishment is *addressed* to the punished person as a creature capable of both speech and understanding (“made in God’s image” and having “reached the age of reason”) and sends the message that he “ought to have known better”. In contrast, imposed cure has no addressee, but a recipient: it is a reaction that could also be appropriate for non-human creatures deprived of reason. From the point of view of its intentions, imposed cure seems to be more humane than punishment, since it does not seek to cause pain; but Lewis argues that, ontologically speaking, from the point of view of its nature, punishment is more humane than imposed cure.

---

<sup>18</sup> Like Duff (2025, 95) we can make a distinction between the primary addressee and the secondary addressee of a response.



However, there is a middle way between Lewis' retributivism and the Clockwork Orange scenario he opposes (in which crime is seen as a disease and punishment is replaced by an imposed cure). One might maintain that responding to wrongdoing is indeed essential if we want to treat the wrongdoer as a person, if we do not want "to give up on him as a moral agent" (Duff 2001, 123), while arguing that such responses need not be punitive in nature. It is for instance the view defended by Narayan in her censure theory: she argues that when we fail to *censure* the wrongdoer and merely react to his wrong (for instance by denouncing it), we treat "the actor as one would treat a natural source of trouble—viz, as an entity not to be directly addressed" (Narayan 1993, 172).

One final remark regarding penal responses. What our analysis reveals is that most definitions of punishment are deficient, as they fail to include the identification component intrinsic to any type of response in the strict sense. This component is not included in the Flew-Benn-Hart definition, i.e. pain intentionally inflicted by the state on an offender for his offense (Hart 2008, 4-5); it is not included in Boonin's definition of legal punishment as authorized intentional reprobative retributive harm (Boonin 2008, 26). Yet this element is necessary if we are to avoid puzzling cases being classified as punishment. Let us consider, for instance, the cases of animal trials in medieval Europe and trials of inanimate objects in Ancient Greece. Will we say that authorized intentional reprobative retributive harm inflicted upon animals and inanimate objects count as punishment for, say, murder? If we are reluctant to say that they are or were punished for murder, we now have a clear explanation: we tend to understand punishment as a response to wrongdoing; now, animals and even more so inanimate objects are unable to perceive harm inflicted upon them after trial as a response to the wrong we attribute to them.

## 5. Strawsonian reactive attitudes and practices

An objection might be raised that I am merely reiterating Strawson's distinction between the objective attitude and the participant attitude; between objective attitudes and reactive attitudes. It could be argued that I am expressing Strawson's view in alternative terms. However, I believe this interpretation is mistaken, and I will clarify why.

First, you may choose to react to a wrong rather than respond to it without adopting an objective stance toward the wrongdoer. Just because you merely react to a crime—for instance by calling the police, as in case 2.B, or by commemorating it—does not mean that you see the perpetrator (or the perpetrators) with an objective eye, "as an object of social policy; as a

subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of” (Strawson 2008, 9).

Second, Strawson contends that “reactive attitudes are essentially natural human reactions to the good or ill will or indifference of others towards us, as displayed in their attitudes and actions” (Strawson 2008, 10-11). Yet this observation does not wholly apply to responses in either interpersonal or criminal justice contexts. You may blame someone who broke a vase out of negligence by telling him that he should have been more thoughtful, though his action shows no ill will, just a lack of reasonable care (Hart 2008, 136). Similarly, strict liability, however unfair it may be, shows that we may be punished with small fines for many actions (e.g. selling expired products through no fault of our own) without being blameworthy. However, the weaker point—the commonplace—that our responses are *influenced* by our beliefs about “the quality of others’ will towards us” (Strawson 2008, 15) is obviously true.

Third, Strawson claims that “reactive attitudes rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard (...) towards ourselves” (Strawson 2008, 15). While this is accurate in some instances, it cannot be generalized to all. When you respond negatively to a personal injury or an action that has hurt your feelings, the underlying demand for regard is easy to see. However, when the criminal justice system responds to the crime of tax evasion with fines or imprisonment, this demand is less easy to perceive. Similarly, some student disciplinary offenses (such as falsifying documents) may require certain responses that are not based on a demand of regard towards anyone—rather, they are based on an expectation of compliance with the rules that has not been met.

Fourth, Strawson (and most of the post-Strawsonian literature) is mainly interested in our human “moral sentiments” (Strawson 2008, 26), whereas I am, as I have emphasized, primarily interested in our behaviors—in what we *do* in reaction or in response to wrongdoing, not in what we *feel* in reaction to wrongdoing. By no means do I wish to deny that what we do and what we feel are closely related. Imagine that you have been wronged and are considering how to respond. You decide to go and see the person who wronged you to explain that you are upset with him, or to show him that you are angry. Even if emotions play a central role in the response you choose, it is still a response in the strict sense (fitting within the C series), in that you are not just brooding over your anger or digesting your resentment: you are confronting the person who wronged you to let him know how you feel.

Fifth, the concept of responses I propose aligns more closely with what Strawson calls the (reactive) “practices” of blame, moral condemnation, and punishment, which he sets aside at the beginning of “Freedom and Resentment” only to reconsider them at the end. One of Strawson’s main points is that we should not lose sight of the fact that, whether we like it or not, these practices are in part expressions of our moral sentiments; they are “not merely devices we calculatingly employ for regulative purposes” (Strawson 2008, 27). This insight applies perfectly to responses to wrongdoing. When we discuss the rationale for a given response (or a type of response), we may put forth its beneficial consequences, such as reducing recidivism. However, this does not mean that this response (or type of response) is a tool designed to produce these beneficial consequences. In fact, it will rarely be perceived as such by the respondent, and even less so by the addressee. What we see in a punishment inflicted on us is primarily the emotion it expresses (e.g. resentment) and the message it conveys (e.g. you “ought to have known better” as C. S. Lewis would say). Even if it is inflicted for our own good, that is not what it immediately means to us.

Sixth, Strawson’s argument against the feasibility of adopting a thoroughly objective attitude applies equally to reactions. It seems, “for us as we are, practically inconceivable” (Strawson 2008, 12) to abandon responses in favor of reactions; to react to harm caused by others exclusively in the same way as we react to damage caused by entities deprived of *logos*; by taking only measures that are not addressed to anyone, that have no dialogical dimension, that do not involve looking each other in the eye. Besides, such a world—which is not ours and never will be—would greatly impoverish human life (Strawson 2008, 14). Most would find this world unappealing, as it would lack the richness of a truly human community.

## 6. An assessment of response retributivism

We are now ready to answer the question that constitutes the title of this article. A response to moral or criminal wrongdoing is a form of action motivated by the wrong. Its defining characteristic is that it is addressed to the wrongdoer, has a confrontational-communicative dimension, and needs to be identified by the wrongdoer as a response to his wrong. Thus, it is perfectly possible to react to a wrong without responding to it. And it seems to me that one of the many meanings of the ambiguous word “responsible” refers precisely to responses in this strict sense: responsible agents are those who are “responseworthy”.

These reflections on the nature of responses to wrongdoing may shed light on the question of what constitutes an *appropriate* response to wrongdoing. For instance, my analysis suggests that letting go or feeling resentment cannot be classified as appropriate responses to wrongdoing; they are not responses to wrongdoing at all, though one could argue that letting go is sometimes the right thing to do, and that resentment may be an appropriate emotion. The question of the appropriate response to wrongdoing, if we use the term “response” in a vague and indeterminate way, is extremely difficult to address. It would probably be easier to address if we made a more rigorous distinction, as I propose, between: first, the question of the appropriate emotion in response to wrongdoing—*is it appropriate to feel pity for the perpetrator of a mass murder in a school?* Second, the question of the appropriate reaction—*is installing metal detectors at schools an appropriate reaction?* Third, the question of the appropriate response (in the strict sense)—*is it appropriate to sentence the perpetrator of a mass murder to an educational measure?* Furthermore, such a distinction would spare us the unproductive search for *the* appropriate response (in the vague and indeterminate sense); it would compel us to acknowledge that, when faced with wrongdoing, several things can be appropriate at the same time: a given emotion, a given security measure, and a given confrontational-communicative act.<sup>19</sup>

This distinction may also help us to evaluate the merits and limitations of response retributivism. With regard to individual non-state actors, the idea that there is a general duty to respond to wrongdoing seems implausible to me. Responses are sometimes too dangerous because of their confrontational nature; in other cases, they are too painful, especially for victims who no longer wish any contact with their aggressor; in other cases, there is no doubt that they will be ineffective; in still other cases, we may lack the standing or authority required to respond to wrongdoing: “X’s mother may have the standing to censure her for stealing cookies out of her cookie-jar, but X’s busybody neighbour does not” (Narayan 1993, 168). The Aristotelian notion that refusing to respond (especially to take revenge) is a sign of servility is simply a reflection of a culture of honor that is no longer ours. Still regarding individual non-state actors, I am not even sure that there is a general duty to react to wrongdoing: sometimes, the best thing to do may be to do nothing. As Brunning and Milam note:

We let go for particular reasons, in response to particular situations of moral conflict, and often with particular aims. And these are often good reasons to do so. When it accomplishes

---

<sup>19</sup> I would like to thank reviewer 2 for bringing this point to my attention.

these aims, letting go can be good (...): therapeutic, liberating, beneficent, and even virtuous. (Brunning and Milam 2022, 15)

That said, even if there is no general duty, I do not wish to deny that in certain specific contexts or situations, individual non-state actors may have a duty to react, or even to respond. The clearest case arises when they bear a “role-responsibility” (Hart 2008). For example, if you are responsible for the safety of a group of children, you have at least a duty to react when one of them is wronged. Likewise, as a parent, it is part of your role to respond to serious wrongdoing committed by your child, provided they are sufficiently mature to understand what you are attempting to communicate.

Now, what about the state? Four cases should be distinguished:

- Response with reaction: The criminal justice system punishes the offender, supplements the sentence with security measures and a treatment order, while the government launches a broad initiative to eliminate or significantly reduce this type of crime.
- Response without reaction: The wrong is minor, and the state considers that a warning is the most appropriate response, requiring no further action.
- Reaction without response: The perpetrator of the murder suffers from serious mental disorders and is unaware of what he has done, rendering him “unfit to stand trial”. In such a case, responding to the wrong he has done would be meaningless (Duff 2025, 94); the person is instead committed to a psychiatric facility.
- Absence of reaction and response: According to the principle of prosecutorial discretion (the “opportunity principle” in French criminal law), a prosecutor may decide not to prosecute an apparent violation of criminal law, since he considers that prosecution would likely do more evil than good (see Dempsey 2009).

As we can see, the state does not have a general duty to react or respond to (legal) wrongdoing, surprising as it may seem. One might object that in *normal* cases, the state has a *pro tanto* duty to respond to (legal) wrongdoing. But such a thesis is too vague to be normatively interesting; it raises more questions than it resolves: What constitutes a normal case? Which response is to be chosen? Which circumstances can defeat or override this *pro tanto* duty? Thus, the truly difficult questions for state actors are rather the *when* question and the *which* question: When to drop charges? When to complement punishment with safety measures, and which ones? When to use an alternative response to punishment, and which one? However, these questions are beyond the scope of this article.

## Acknowledgments

This first version of this article was written as part of the REACT project, funded by the Agence Nationale de la Recherche (ANR-20-CE28-0012). I presented it at the German-French-Italian Workshop “Themes on Moral Psychology” at LMU Munich in 2023. I am grateful to the participants, and in particular to Jörg Löschke, for their precious comments. Finally, I would like to thank the two anonymous reviewers for their helpful suggestions, which contributed to improving the article.

## REFERENCES

- Anscombe, G. E. M. 2000. *Intention*. Cambridge: Harvard University Press.
- Aristotle. 2007. *On Rhetoric: A Theory of Civic Discourse*. Translated by George A. Kennedy. New York: Oxford University Press.
- Boonin, David. 2008. *The Problem of Punishment*. Cambridge: Cambridge University Press.
- Brunning, Luke, and Per-Erik Milam. 2022. “Letting Go of Blame.” *Philosophy and Phenomenological Research* 106: 720-740. <https://doi.org/10.1111/phpr.12899>.
- Copp, David. 1997. “Defending the Principle of Alternate Possibilities: Blameworthiness and Moral Responsibility.” *Noûs* 31 (4): 441-56. <https://doi.org/10.1111/0029-4624.00055>
- Dahan Katz, Leora. 2020. “Response Retributivism: Defending the Duty to Punish.” *Law and Philosophy* 40 (6): 585–615. <https://doi.org/10.1007/s10982-020-09386-3>
- Duff, Antony. 2025. “Communicative Theory.” In *The Oxford Handbook of the Philosophy of Punishment*, edited by Jesper Ryberg, 90-112. Oxford: Oxford University Press.
- Duff, Antony. 2001. *Punishment, Communication, and Community*. Oxford: Oxford University Press.
- Durkheim, Émile. 2013. *The Division of Labour in Society*. Translated by W. D. Halls. Basingstoke: Palgrave Macmillan.
- Feinberg, Joel. 1965. “The Expressive Function of Punishment.” *The Monist* 49 (3): 397–423. <https://doi.org/10.5840/monist196549326>
- Hampton, Jean. 1988. “The Retributive Idea.” In *Forgiveness and Mercy*, by Jeffrie Murphy and Jean Hampton, 111-161. Cambridge: Cambridge University Press.
- Hart, H. L. A. 2008. *Punishment and Responsibility: Essays in the Philosophy of Law*. Oxford: Oxford University Press.

- Hieronymi, Pamela. 2001. "Articulating an Uncompromising Forgiveness." *Philosophy and Phenomenological Research* 62 (3): 529-55. <https://doi.org/10.1111/j.1933-1592.2001.tb00073.x>.
- Holmgren, Margaret Reed. 2012. *Forgiveness and Retribution: Responding to Wrongdoing*. New York: Cambridge University Press.
- Kant, Immanuel. 1991. *The Metaphysics of Morals*. Cambridge: Cambridge University Press.
- Lewis, C.S. 1987. "The Humanitarian Theory of Punishment." *Issues in Religion and Psychotherapy* 13 (1): 147-53.
- Macnamara, Coleen. 2015. "Reactive Attitudes as Communicative Entities." *Philosophy and Phenomenological Research* 90 (3): 546-69. <https://doi.org/10.1111/phpr.12075>.
- McKenna, Michael. 2012. *Conversation and Responsibility*. New York: Oxford University Press.
- Metz, Thaddeus. 2008. "The Nature of Reactive Practices: Exploring Strawson's Expressivism." *South African Journal of Philosophy* 27 (3): 227-41. <http://dx.doi.org/10.4314/sajpem.v27i3.31514>
- Murphy, Jeffrie. 1988. "Forgiveness and Resentment." In *Forgiveness and Mercy*, by Jeffrie Murphy and Jean Hampton, 14-34. Cambridge: Cambridge University Press.
- Narayan, Uma. 1993. "Appropriate Responses and Preventive Benefits: Justifying Censure and Hard Treatment in Legal Punishment." *Oxford Journal of Legal Studies* 13 (2): 166-82. <https://doi.org/10.1093/ojls/13.2.166>
- Nussbaum, Martha C. 2016. *Anger and Forgiveness: Resentment, Generosity, Justice*. New York: Oxford University Press.
- Seneca. 2012. *Anger, Mercy, Revenge*. Translated by Robert A. Kaster and Martha C. Nussbaum. Chicago: University of Chicago Press.
- Shoemaker, David. 2015. *Responsibility from the Margins*. Oxford: Oxford University Press.
- Strawson, Peter F. 2008. *Freedom and Resentment and Other Essays*. Abingdon: Routledge.
- Waldron, Jeremy. 2010. "Inhuman and Degrading Treatment: The Words Themselves." *Canadian Journal of Law & Jurisprudence* 23 (2): 269-86. <https://doi.org/10.1017/S0841820900004938>.
- Walker, Margaret Urban. 2012. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. Cambridge: Cambridge University Press.
- Watson, Gary. 2019. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In *Perspectives on Moral Responsibility*, edited by John Martin Fischer and Mark Ravizza, 119-48. Ithaca: Cornell University Press.

