Philosophical Studies (2024) 181:2219–2246
https://doi.org/10.1007/s11098-023-02077-4

Difference-making and the control relation that grounds responsibility in hierarchical groups

Johannes Himmelreich[1]

Accepted: 3 November 2023 / Published online: 8 December 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
Hierarchical groups shape social, political, and personal life. This paper concerns the question of how individuals within such groups can be responsible. The paper explores how individual responsibility can be partially grounded in difference-making. The paper concentrates on the control condition of responsibility and takes into view three distinct phenomena of responsibility in hierarchical groups. First, a superior can be responsible for outcomes that her subordinates bring about. Second, a subordinate can be responsible although she is unable to prevent the outcome she brings about. Third, a superior can sometimes be responsible to a greater degree than her subordinates. It is argued that difference-making, as an interpretation of the control condition that partially grounds responsibility, accounts for all three of these phenomena within a limited but significant range of circumstances and can hence partially ground individual moral responsibility in hierarchical groups. The paper provides an element of a theory of individual responsibility to complement theories of corporate responsibility.

Keywords Moral responsibility · Collective responsibility · Difference-making · Causation · Organizational ethics · Social ontology · Respondeat superior

1 Introduction

In early February 1989, rumors raged the streets of East Berlin that the *Schießbefehl* had been suspended—the order to shoot anyone trying to cross the border. Seeing his chance, twenty-year-old Chris Gueffroy approached the fences of the Berlin Wall. But the rumors were just that. Gueffroy became the 67th and last person to be shot at the Berlin Wall (Hertle and Nooke 2009, 18–25). Gueffroy was killed by a hierarchi-

✉ Johannes Himmelreich
jrhimmel@syr.edu

[1] Syracuse University, 900 South Crouse Ave Syracuse, Syracuse, NY 13244-1020, USA

cal group. Three years later, GDR soldiers as well as party leaders were convicted of manslaughter and murder.

Many groups—businesses, government agencies, or military organizations—are hierarchical.[1] Hierarchical groups lay off employees, deport immigrants, and kill in wars. This paper concerns the question of how individuals in such groups can be responsible.[2] I argue that *difference-making*, as defined by List and Menzies (2009), provides a promising partial ground of such individual responsibility: Individual responsibility is grounded in the difference individuals make.

Understanding responsibility in hierarchical groups is relevant for business, organizational, and military ethics. It is crucial for issues in just war theory—such as the permissibility of retributive operations, or for post-war justice and reconciliation. Moreover, accounts of responsibility in hierarchical groups may support dogmatic development in criminal law.[3] Hierarchical groups are a common topic in legal philosophy (Roxin, 1963, 2011; Smidt, 2000; Kutz, 2000, Chap. 5; Danner and Martinez 2005; May, 2005; Weigend, 2011; Eldar, 2013; DeFalco, 2013). In moral philosophy, debates on collective responsibility have instead concentrated on joint actions, pairs of individuals, or the corporate responsibility of a group as a whole (French, 1979; Pettit, 2007; List and Pettit 2011; Isaacs, 2011; Bazargan-Forward, 2017).[4] This paper provides metaphysical foundations for extant accounts in philosophy of law, it complements accounts in philosophy by exploring the control condition of responsibility and couching the proposal in an alternative causal framework, and it emphasizes a role for individual responsibility in certain collective contexts.[5]

---

[1] The debate over whether there are groups and whether they are agents is irrelevant for present purposes (since there might be an *individual* at the top of the hierarchy—hence the action of a hierarchical group would be, in a way, the action of this individual, which is carried out by individuals down the "chain of command").

[2] To understand what it means for a (hierarchical) group to act, I follow the two-level picture of Isaacs (2011): When a group acts there are two actions (or facts pertaining to two levels)—the action of an individual acting on behalf of the group, and the collective action, which the individual brings about (in some sense) but of which she is not the agent. My view differs from Isaacs' in that I account for this difference via conditionals instead of what descriptions of the actions are possible (see Sect. 3). Moreover, to be applicable to hierarchical groups with many levels, I extend Isaacs' account from a two-level to a multi-level view (this seems consistent with her view). Schematically, when a hierarchical group acts, for each level on a hierarchy, there is an individual action to be found. When there is no individual at the top of the hierarchy and the group is a corporate agent (cf. List and Pettit 2011, or Isaacs, 2011), then the individual actions together bring about, but are different from, the action of this group as a whole. I understand actions, of individual or corporate agents, as events that are caused in the right way by an agent's intentions.

[3] The International Criminal Court (ICC) is equipped with an assemblage of various modes of collective responsibility from different legal traditions. In recent cases the court was split on the question of how these distinct approaches combine and interact and what distinguishes a joint criminal enterprise from indirect co-perpetration (Manacorda and Meloni 2011; Vest, 2014).

[4] Nevertheless, there is a significant existing and growing literature on hierarchical groups (Feinberg 1970b, 227–28; Wasserstrom, 1980; Walzer, 2004; Crawford, 2007; Isaacs, 2011, Chap. 4; Shapiro, 2014; Bazargan-Forward, 2022).

[5] Similarly, Joel Feinberg (1970b) draws a distinction between collective responsibility and responsibility in hierarchical groups. Kutz (2000, 11) likewise criticizes the "paradigm" of collective responsibility as "unsuited to the depersonalized, hierarchic, bureaucratic, but nonetheless collective institutions that characterize modern life."

The paper is structured as follows. In Sect. 2, I start with a motivating puzzle, characterize hierarchical groups, and delineate the topic. I then discuss existing solutions to this puzzle in Sect. 3. Next, in Sect. 4, I build on the characteristics of hierarchical groups from Sect. 2 to offer an account of how difference-making can partially ground individual moral responsibility. I discuss limitations and implications of this account in Sect. 5.

To avoid confusion: By "responsibility" and its cognates I mean retrospective moral responsibility. Philosophers have put forth a good variety of analyses of responsibility (Watson, 1996; Smith, 2006; Shoemaker, 2011). But in this paper, I am concerned with what *grounds* responsibility (the grounding question) rather than with what responsibility *is* (the analytical question).[6] What I say should be compatible with various existing analyses of responsibility.[7] In particular, my account is compatible with understanding responsibility as praise- or blameworthiness.

This paper is on one partial ground of responsibility: the control condition (in contrast to the epistemic condition, or other conditions that there may be).[8] We are not responsible for everything we do. We are responsible at most for some aspects and consequences of our conduct—things that we have done or that we refrained from doing. I will use "agency" as the name of some such relation that, put metaphorically, circumscribes responsibility's *reach* by identifying those things for which someone *might* be responsible, and that hence partially grounds moral responsibility.[9] As Rosen (2010) remarks: For *which* consequences agents are (not) responsible is a question "about which philosophers have been strangely silent."[10] This agency relation—that someone *did* something, is the agent *of* something, or exercised agency *with respect to* some consequence—is only a partial ground of moral responsibility in two respects. First, the agency relation is only *one* ground of responsibility among several others (such as those captured by the epistemic condition). Second, some argue that the agency relation grounds responsibility only *sometimes* (that someone is an agent of something is neither necessary nor sufficient for their responsibility).[11]

---

[6] This distinction is due to Rosen (2015) and I follow the understanding of "ground" presented there.

[7] With one limitation: The understanding of "responsibility" in moral philosophy (which I follow), and the role the concept plays here, differs from the understanding in the literature on just war theory.

[8] I use "agency" and "control" largely interchangeably. I prefer to talk of "agency" rather than "causation" because, whatever the agency condition may be, it is a further question how agency relates to causation (one could, for example, deny transitivity of one but not the other).

[9] I focus on agency as a *relation* (between agents and outcomes) in contrast to agency as a *predicate* (i.e., an entity capable of acting). Agency as a relation seems central at least for some moral responses. In the terms of Shoemaker (2015), the kind relevant here would be "responsibility as accountability." There are limits to this conception (see e.g. Smith, 2006).

[10] Meanwhile, the agency relation has received sustained attention in what is known as the *control problem* in the literature on group agency (Pettit, 2007; Searle, 2010, 50–55; List and Pettit 2011, 160–63; Szigeti, 2014; Roth, 2014).

[11] Consider, relatedly, how causation by itself can hardly ground moral responsibility. (1) Agents are not responsible for *all* causal consequences of their actions (Rosen, 2010). (2) Agents might be responsible for things they did not cause (Sartorio, 2004). Moreover, some analyses, or forms, of responsibility do not require agency with respect to some event or outcome.

2 Hierarchical groups

The case of Chris Gueffroy illustrates hierarchical groups in practice. But hierarchical groups are also theoretically interesting. Consider a simple case: A commander has several soldiers under her command. She orders these soldiers to shoot an innocent victim. And they comply. Each soldier goes to their station; the soldier whose station the victim passes first shoots the victim. How can we explain everyone's responsibility?

It seems natural to explain everyone's individual responsibility in terms of their causal contribution, their control, or more broadly, their agency. But doing so is not as straight-forward as it may seem. There are two families of theories of causation: production theories and dependence theories (cf. Hall, 2004). Roughly, the former see causes as mechanisms that produce their effect through a local transfer of some force, such as billiard balls hitting each other or a lightning setting a tree on fire. The latter theories instead see causes in terms of necessary and sufficient conditions on which the effects depend. Sometimes these theories arrive at different conclusions—the presence of oxygen and the absence of rain are causes of the tree catching fire on dependence theories, but not on production theories. In this case of a commanded killing, however, both kinds of theories are hard pressed to explain everyone's responsibility.

When we turn to production theories of causation, such as "such as mark-transfer theories and conserved-quantity-transfer theories," then, Bazargan-Forward (2017, 111) observes, "what [a superior] does might not count as causally contributing to [a subordinate's] wrongful act."[12] Production theories of causation are hence hard pressed to explain why the commander is responsible.

Dependence theories of causation, by contrast, have easy answers for why the commander is responsible. Since the commander is a "*sine qua non* … [this] provides a prima facie basis for regarding [the commander] as at least somewhat morally responsible" (Bazargan-Forward, 2017, 111). But dependence theories are instead hard pressed to explain why subordinates are responsible. For example, the victim's death does not counterfactually depend on any subordinate's conduct. It is a case of redundant causation: If one subordinate had resisted the order, another solider would have shot the victim instead. Thus, at least on simple counterfactual dependence, no soldier is an agent of the victim's death (cf. Lawson, 2011, 240).

To make matters worse, if there are more than two levels in the hierarchy, the victim's death does not counterfactually depend on our commander's order either. If there is a commander above our commander and other commanders left and right, our commander's conduct is also redundant: If she had refused to give the order to shoot the victim, another commander would have. Hierarchical groups have such redundancies on almost every level.

---

[12] Bazargan-Forward (2017) examines non-hierarchical pairs of individual agents and thus does not speak of "superior" or "subordinate" but of an enabling and an enabled individual "P1" and "P2" respectively. But the puzzle for grounding the control condition of responsibility in a causal process theory is sufficiently similar.

In sum, whether and how agency grounds responsibility in hierarchical groups is considered somewhat of a puzzle in the literature. This is not to suggest that there is a yet unsolved or overlooked theoretical problem to which I offer the first solution.[13] To the contrary. Theories of causation generally, and the debate on collective responsibility specifically, already offer relevant solutions (I discuss some below). This paper applies one theory from the former to contribute an alternative account to the latter.

2.1 Characteristics

In this simple case as well as in the case of Chris Gueffroy, hierarchical groups exhibit three relevant characteristics: asymmetry, control, and redundancy (cf. Buchanan, 1996; Prosecutor v. Katanga and Chui 2008).

First, hierarchical groups have asymmetric authority relations (formally, a partial order): Some are superiors to subordinates (or they occupy a "higher level"), and they can direct their subordinates, but not usually the other way around.[14] This hierarchy enables a *vertical division of labor*: Superiors give directives that are typically incomplete, or at least not maximally specific.[15] Hence, subordinates can—within limits—carry them out in more than one way.[16] By contrast, many existing accounts of shared agency assume forms of equal or symmetric cooperation, or a *horizontal division of labor* as seen in sports teams, mobs, committees, or juries (e.g. Bratman, 2014, 7, 85).

Second, along the lines of this asymmetric structure, we find a functional dependence. Subordinates not only *ought* to respond to their superiors, but they *in fact* generally do respond to superiors in certain ways. This is because individuals higher up have some degree of control or influence over their subordinates.[17]

Third, hierarchical groups often exhibit a certain degree of redundancy to ensure that the directives of superiors are carried out. Typically, redundancy in hierarchical groups is achieved by associating several subordinates with each superior. In result, the structure of a hierarchical group, of the kind considered here, looks like a pyramid or an inverted tree. At the top is one node from which subordinate nodes spring that, in turn, serve as superiors to further subordinates, and so forth.[18]

---

[13] Perhaps the phenomenon of increasing responsibility (introduced below) received somewhat less attention and still benefits from clarification (e.g., how should the expression of someone being "more responsible" be understood?) before it can be explained.

[14] These superiors might have superiors in turn. Thus, an individual can be both a superior (with respect to some) and a subordinate (with respect to others). I assume that the authority relation on the group is well-ordered.

[15] "Directives" is sufficiently general to cover not only "orders" but also "decisions," "plans," or "policies," which play a similar role. I am grateful to a reviewer for suggesting this term.

[16] This is a phenomenon well-known in law (DeMott, 2014), public administration (Lipsky, 1980; Heath, 2020, Chap. 6), and principal–agent theory (cf. Buchanan, 1996). Sunstein and Ullmann-Margalit (1999) investigate the general class of second-order decisions, of which such delegation is but one type (See also Shapiro, 2014).

[17] "Control" might be misleading since superiors' influence is not always intentional. Whenever a superior issues a directive, there is a risk that the directive might be misunderstood or that it may have unforeseen consequences. Furthermore, superiors do not always exert control explicitly; subordinates might understand the intentions of their superiors tacitly.

2.2 Responsibility

Hierarchical groups exhibit three phenomena relating to responsibility. Not only are, first, superiors responsible but so are, second, subordinates. In the case of the Berlin Wall shootings, those up the chain of command who issued the order to shoot are responsible, as well as the soldiers who carried out the order.[19] Third, at least in some cases, responsibility can be increasing; that is, superiors can, in some sense, be *more responsible* than their subordinates.[20] "Can be" because responsibility may increase in some groups and in some cases but not in others.[21]

Many defend increasing responsibility (cf. May, 1992, 123; 2005, 154; Isaacs, 2011, 115). And the phenomenon can also be seen in criminal law: The commanders in the Berlin Wall shootings, for example, received longer sentences. For an even clearer statement, consider the case of Adolf Eichmann. When a court in Jerusalem sentenced Eichmann to death in December 1961, it recognized that in hierarchical groups, responsibility tends to increase with hierarchy.

> [T]he extent to which any one of the many criminals were close to, or remote from, the actual killer of the victim, means nothing as far as the measure of his responsibility is concerned. On the contrary, in general, the degree of responsibility increases as we draw further away from the man who uses the fatal instrument with his own hands and reach the higher ranks of command (The Attorney General v. Eichmann, 1961, 197).

Increasing responsibility can be seen also in private organizations. For example, the vice-president of a corporation that produces landmines is, in some sense, more responsible than its shipping clerk (Kutz, 2000, 158; Lawson, 2011, 240).

All this suggests that the three characteristics and three phenomena delineate a domain of hierarchal groups that is broader than military organizations and extends at least to some private corporations. However, an account developed for this domain does not extend to unorganized groups (which might have a structure, albeit of a different kind), or mobs, who lack this combination of asymmetric structure, function, and organized redundancy.

---

[18] It should be noted that this introduces an important further assumption, namely, that each subordinate is subordinate to exactly one superior. Redundancy may not be an essential feature of hierarchical groups, but it is an essential feature of certain organizations. Since such organizations—be it military, bureaucratic, or corporate—are central to social, private, and political life, it is organizations and hence hierarchical groups with redundancy, on which I focus here.

[19] The description is ambiguous about the objects of responsibility, that is, what superiors and subordinates are responsible *for*. I understand them as individuals who are responsible for individual actions or outcomes.

[20] This could be spelled out by saying that they are responsible *to a greater degree* (for the same action as for which their subordinates are responsible), or by saying that they are responsible *for a worse action*. Since I leave issues of action individuation aside for now, the two can be used interchangeably.

[21] To be clear: Superior and subordinate responsibility, the first two phenomena, may likewise be characteristic but not essential. I do not claim that superiors and subordinates are responsible in *all* hierarchical groups or in *all* cases.

3 Existing work

An account of responsibility in hierarchical groups needs to make sense of the three phenomena. It should explain the individual responsibility of superiors and subordinates and explain increasing responsibility. Concentrating on the agency relation to ground moral responsibility, a central challenge is to give an account that is unified (in that it is characterized by the same criteria for all individuals in a hierarchy). Moreover, an account should distinguish individual contributions. A superior and a subordinate seem responsible on somewhat different grounds: A commander, for example, is responsible, in part, because she gave an order. A soldier is responsible because they carried out or did not resist this order. Finally, an account should scale, that is, it should be applicable to hierarchical groups with many levels. Meeting these desiderata is a theoretical challenge—after all, as more levels of hierarchy are involved, the initial intuitive clarity of talking of "individual contribution" is soon lost.

Several authors have tackled this challenge. Christopher Kutz, Larry May, as well as Tracy Isaacs, and Matthew Braham and Martin van Hees put forward accounts that could be applied to the three phenomena (Kutz, 2000, Chap. 5; May, 2005, Chap. 8; Isaacs, 2011, Chap. 4; Braham and van Hees 2009; 2012). But each of them either concentrates on agents' intentions and the epistemological condition of responsibility not on individual agency as a ground of individual responsibility (Kutz and May), or their account doesn't extend to hierarchical groups (Braham and van Hees), or it can't clearly distinguish the different contributions or outcomes associated with individuals on different levels in the hierarchy (Isaacs).

Kutz explains increasing responsibility with facts about agents' intentions.[22] A superior is more responsible than a subordinate because a superior has an "executive *intention*" whereas a subordinate only has a "subsidiary *intention*" (Kutz, 2000, 160 my emphasis). Both *participate* in the action of the group—this the limited extent to which the control condition comes in—but the difference in their responsibility is explained by a difference in their intention: by "the attitudes they take towards the success of the activity" (Kutz, 2000, 159).[23] In fact, Kutz explicitly opposes explaining increasing responsibility as having to do with agency (2000, 52), and others agree (Lawson, 2011, 240–42). Isaacs (2011, 118) observes similarly that "Kutz detaches individual responsibility from causal contribution." Thus, Kutz' proposed answer to the puzzle of who is responsible for a commanded killing doesn't rely on individual agency at all. It concentrates almost entirely on individuals' intentions and doesn't seek to incorporate the intuitive idea that someone's responsibility depends on what they did.

---

[22] Kutz (2000, 52) dismisses appealing to "degrees of causation" because "independent normative interests are doing the real work."

[23] Specifically, the superior "must identify … and so must associate herself directly with [the] morally relevant characteristics" of the action (Kutz, 2000, 159).

May, like Kutz, concentrates on agents' intentions. He investigates subordinate responsibility primarily with an eye to intent (May, 2005, 125–27); defends "a special intent requirement" for superior responsibility (2005, 139); and understands hierarchical groups generally in analogy with conspiracies (2005, 143). Occasionally, May relates superior and subordinate responsibility to the control condition (cf. 1987, 89–106; 2005, 139–48); contending that superiors' responsibility is grounded in negligent omissions (2005, 142).

My account differs from—and improves on—May's in two main ways. First, I don't think superiors are responsible mainly because of their negligence or their omissions. Superiors are responsible for giving directives at least much as they are for failing to act. My account captures this active agential role of superiors (and subordinates). A second difference between May's proposed solution and my account is that May seeks to hold individuals responsible for *collective* actions.[24] After all, May's overall interest is in explaining why heads of state are responsible for crimes against humanity, such as genocides, seen as collective actions. My account, instead, concentrates on individual responsibility for individual actions and spells out how superiors and subordinates differently contribute to some outcome.

Likewise does the account developed by Tracy Isaacs. Isaacs concentrates on individual responsibility for individual actions in hierarchical groups, and she seeks to explain the three phenomena via the control condition of responsibility (2011, Chap. 4). Her account takes as a theoretical primitive that for each basic action there is a set of action descriptions, such that any individual action can be re-described in terms of any of its causal consequences (2011, 100–101).[25] Isaacs uses this to explain how individuals are responsible *in*, even if not *for*, collective actions: A superior is individually responsible *in* the collective action of killing an innocent victim because her action of ordering to kill this victim can be redescribed as this victim being killed by a subordinate.

But this approach has two main problems. First, it is not clear how it distinguishes superiors' and subordinates' differential contribution to some outcome. Both individual actions can, likewise, be described in terms of the same causal consequence: the victim's death. Second, and more fundamentally, it is unclear how the set of action descriptions for a given action is to be delimited. Without further restrictions on how actions can be described, an agent would be an agent of too many consequences (I discuss how my account addresses such a problem towards the end of Sect. 4.1). But we need to limit what individuals are agents of. To avoid circularity, any account based on action descriptions needs a certain theory of causation or an independent theory of action individuation (see Bratman, 2006).[26]

My account rests on theoretical underpinnings that avoid these problems. My account offers explicit conditions to identify—and limit—what individuals are agents of (it moreover explicates some assumptions about action individuation, in Sect. 4.3). In this way, my account develops what an account that starts with action descriptions presupposes.

---

[24] Isaacs (2011, 105) supports this reading.

[25] Isaacs (2011, 101) draws on the so-called Accordion Effect due to Feinberg (1970a).

[26] Isaacs (2011, 106) seems to recognize the problem and ends up rejecting the so-called Accordion Effect—that actions remain the token-same under different descriptions.

Finally, Braham and van Hees (2009; 2012) develop an account of individual responsibility in collective settings that likewise centers on the agential condition. Their account of agency is based on the NESS test—another way of understanding "difference making"—which they formalize using game theory.[27] However, their account does not extend to hierarchical groups. Individuals in hierarchical groups "play" conditional strategies: Subordinates perform certain actions, *if* so directed by a superior. But the NESS test cannot deal with such conditional intentions or strategies (Braham and Hees 2009, 615).[28] Moreover, the account that I develop is based on the machinery of possible-world semantics, which, at least to some, might be a preferable theoretical approach (because of its greater generality, simplicity, or some other theoretical virtue). For instance, my account has resources to address the so-called Queen of England problem (e.g. Menzies, 2011; Himmelreich, 2023). Analogous solutions for Braham's and van Hees' account have not yet been developed (2009, n. 7).

4 Agency and responsibility in hierarchical groups

To explain responsibility in hierarchical groups, let's expand on the earlier motivating case by adding some details: Anne, a military commander, directs her team to shoot an innocent victim, Collin. The team consists of several soldiers, Bert, Ben, Bob, and some others. Each is dispatched to a post and each of them is sufficiently motivated to follow Anne's directives, although each may comply for different reasons. Bert's post is the one that Collin passes first. Bert knows that if he were not to shoot, then someone else would. Bert also believes with good reason that he would be faced with mildly severe sanctions if he disobeyed Anne's order. Bert goes ahead and shoots Collin.

Yet, the description of this case still leaves some things open. It does not specify how many soldiers are in Anne's team, what motivates the soldiers, and what sanctions are in store for Bert. But for present purposes, these aspects do not matter very much. What matters is that Anne gives a command to kill an innocent victim (by assumption, a moral wrong). If Anne gives the command, then *someone* will shoot and kill Collin. Collin would die even if Bert were not to shoot.[29] Moreover, in contrast to subordinates operating some murderous machine that requires each to do their part to bring about Collin's death at all, this case involves no joint action or horizontal division of labor but only *vertical* division of labor (at least as far as the act of shooting Collin is concerned). In other words, Bert or any of the other soldiers would shoot Collin individually, they don't shoot Collin together. However, neither Anne nor any of her soldiers have a fully exculpating excuse. Concerning the looming sanctions and the soldiers' motivations, it suffices to assume that the sanctions for anyone refusing to follow directives are mildly severe. That is, the sanctions are sufficiently harsh to dispose subordinates' compliance and yet sufficiently bearable to not excuse them for their conduct. Moreover, the model case is silent about the group's normative context. Bert's responsibility overall depends on the normative structure within which the group operates; in particular, on Anne's authority and hence whether her directives give soldiers a reason to act (cf. Estlund, 2007; Crawford, 2007). But the issue here is not *whether* someone is responsible but only one aspect of *why* they are when they are. Finally, this model case does not clearly exhibit increasing responsibility (I return to this phenomenon later). In sum, even though this model case above involves vast simplifications and leaves many things open, it should resemble paradigmatic cases, like the Berlin Wall shootings, in crucial respects.

---

[27] The NESS test defines c as a cause of e if and only if c is a member of a set of conditions such that these conditions (i) are sufficient for e, (ii) obtain, and (iii) c is necessary for the sufficiency of this set. See Braham and van Hees (2009, 613).

[28] Braham and van Hees (2012, 615) state: "the NESS test has problems dealing with some types of strategies, namely, those comprising conditional actions." This is irrelevant, for example in Frankfurt cases, where some (i.e., Black) have such conditional intentions, if the agents whose responsibility we seek to understand (i.e., Jones) do not have such intentions. But this is not the case for hierarchical groups.

[29] For now, let me assume this local determinism for simplicity (and in line with the characteristic of functional dependence stated above). I relax this assumption to discuss the possibility of a collective resistance against Anne's command later. That *someone* will shoot and that there is more than one shooter, is the assumption, in terms of List and Menzies (2009, 496), that Anne's directive is *realization-insensitive*,

The account that I put forward consists of two parts.[30] The first part consists of dependence conditions to spell out the idea that individuals make a difference to some outcome. This part of the account is stated in terms of two conditionals. The second part concerns the individuation of outcomes. I formalize the intuitive idea that superiors control a course of events only in broad strokes. Subordinates, in comparison, exercise more fine-grained control over what happens, given that the coarse-grained direction has already been fixed. Accordingly, the hierarchy of individuals is associated with a hierarchy of outcomes. This second part comes in to explain subordinates' responsibility as well as to scale up the account (roughly: just as there is someone in case Bert fails to carry out Anne's directives, there might be someone to play Anne's role in case she fails to carry her superior's directives).

4.1 Superiors' responsibility

To explain superiors' responsibility, we only need the following first part of the account. Consider again the case of Anne ordering Collin to be shot by someone. Two things seem to be the case. First, had Anne *not* given a directive to shoot Collin, then he would not have been shot.[31] Furthermore, if Anne *were* to give the directive to shoot Collin—that is, even if the situation were slightly different in some ways but Anne were to give the directive to shoot Collin—then Collin would still be

---

that is, the directive is carried out (Collin is shot) even under small perturbations of how this directive is realized.

[30] For both parts, I draw on an account of difference-making that was put forward in the context of mental causation (List and Menzies 2009).

[31] For reasons discussed below, this is not quite correct. Instead, the result is better described as "then he would not have been shot *on Anne's order*" (similar modifications will need to be made for the case of Bert). This is to avoid the problem of redundant causation, following in some ways a discussion of van Inwagen (1983, 173–74).

shot. This is an instance of two conditions that Lewis (1973b) initially calls "causal dependence"—I will refer to them as the *negative* and *positive conditional* respectively. The negative conditional captures that had Anne not given the order, then Collin would not have been shot (at least not on her order). The positive conditional captures that Anne can ensure that Collin is shot. In a way, each of the two conditionals represents a different aspect of agency. The positive conditional represents the idea of *agency as implementation*, as an individual seeing to it that something will be the case. The negative conditional, by contrast, represents the idea of *agency as prevention*.

The negative conditional is a subjunctive counterfactual conditional. We can evaluate whether it is true using standard semantics for counterfactuals, such as that by Lewis (1973a). This semantics builds on the construction of representing different *possibilia*, such as possible worlds, ordered by increasing distance to the actual world. The negative counterfactual for the given case of Anne is true if in all nearby worlds in which Anne does *not* give the directive to shoot Collin, it is the case that Collin is *not* shot.[32]

The positive conditional should also be read as a subjunctive conditional, but it is not a counterfactual conditional. The statement in the antecedent—that Anne gives the order—is actually true. The idea behind reading this conditional as a subjunctive conditional is that we want to examine whether the outcome of Collin being shot depends on Anne's directive robustly. Will Collin be shot in *all* relevantly similar situations in which Anne gives the order? To capture this robust dependence between Anne giving the directive and the outcome of Collin being shot, we can still rely on a kind of counterfactual reasoning: We ask the question of whether these two things, the directing and the shooting, co-occur reliably under changes in circumstances. To answer this question, the semantics for counterfactuals can be used, if the semantics is suitably amended. The amendment needed is that *other* possibilities may be as close to the actual world as the actual world is to itself (Lewis, 1973a, 29). In effect, conditional statements with true antecedents—such as the positive conditional—are evaluated relative to a set of nearby possible worlds that may contain *more* possibilities than only the actual world.[33]

Agency in hierarchical groups is defined by these two conditions—the conditions of "difference-making" as I propose to understand it here. An individual *makes a difference* to some outcome if and only if, given the amended semantics, the positive and negative conditionals are true. Anne is responsible for Collin's death (partly) in virtue of her making a difference to it. This is an example of how superiors' agency grounds superiors' responsibility.

To state this definition of agency in full generality, let a be an individual—in this case Anne—and let i be an intention of a.[34] Alternatively, it could be said that i stands for a directive that *a* gives. But if the proposal should be generalized beyond contexts of directives, *i* needs to be an intention or an intention-like mental state of *a*. Finally, let *x* be some outcome that actually occurs.[35] Agency in hierarchical groups is then defined as the following dependence. The individual *a* makes a difference to *x* if and only if, if *a* were *not* to have *i*, then *x* would *not* occur; conversely, if *a* were to have *i* then *x* would occur. Agency, or difference-making, with respect to an outcome is defined as this dependence of an outcome on an agent's intention.

---

[32] As so often with semantics for counterfactual conditionals, what counts as nearby must be taken as given. Furthermore, the counterfactual is vacuously true if there is no possible world in which Anne does not give the directive to shoot Collin. For brevity, I set aside vacuous truth.

[33] The so-called strong centering assumption is replaced with the alternative weak centering assumption (Lewis, 1973a, 26–29; List and Menzies 2009).

[34] An intention here is any mental entity that plays a certain functional role. It may be a belief–desire pair, an intention-in-action, or a proximal intention.

This combination of negative and positive conditional should be familiar to many, since this is how Lewis (1973b, 563) initially defines causal dependence, before setting the positive conditional aside. This raises at least two questions. First, what is the purpose of the first, the positive, conditional? Second, how are well-known under-generation problems with the second, the negative, conditional avoided?

First, the role of the positive conditional is to limit individuals' agency. If agency were only defined by the negative conditional, then any agent who makes a difference to one thing would be the agent of far too many things. Lewis (1987, 184–88) discusses the case of writing a reference letter for a student. With only the negative conditional, by writing a strong letter for a student, Lewis would make a difference not only to his student getting the job, but also to several other candidates not getting the job, to them pursuing different careers, to them meeting their partners, to their children being born and passing away. This may or may not be an acceptable implication for a theory of causation, it is not for a theory of agency.[36] Lewis thus suggests that the relevant relation needs to be one that is *insensitive* to circumstances. The positive conditional, evaluated within the amended semantics, is one way of spelling out Lewis' suggestion (see Woodward, 2006 for a related proposal).[37]

Of course, one might argue, insofar as we are ultimately interested in responsibility (and not just agency), superiors' responsibility could be limited in a different way than by limiting their agency. Other conditions of responsibility, and specifically the epistemic condition, could be used to limit superiors' responsibility (cf. Rosen, 2010, 433) and to avoid this over-generation problem that motivates the positive conditional. In this sense, the positive conditional seems under-motivated.[38]

The epistemic condition certainly is *one* limit on an agent's responsibility (and likewise a partial ground of responsibility)—ignorance sometimes excuses. But the positive conditional has a role to play insofar as the epistemic condition can't limit responsibility in *all* relevant cases.[39] An agent may foresee consequences of their action without being responsible for them. Edward the knife sharpener sharpens Giulia's kitchen knives. He foresees that she might cut herself. She does. He's the only sharpener in town. Had Edward not sharpened the knife, Giulia wouldn't have cut herself. But Edward is not responsible. This or similar cases of "intervening agency" suggest limiting responsibility also via agency. After all, the point of partially grounding responsibility in agency is to account for the idea that an agent is responsible at most for consequences of her own conduct, or outcomes that are "glued" to her in some way (Rosen, 2010, 422). The positive conditional allows us to say that Edward is *not* an agent of Giulia's injury, but that Anne *is* an agent of Collin's death;[40] or that Lewis is an agent of writing a reference letter, but *not* of some rejected candidate meeting their partner.

---

[35] Outcomes are represented by sets of *possibilia*, such as possible worlds. I say more about the individuation of outcomes later.

[36] Two independent reasons for this. First, it seems to violate too crassly pre-theoretic judgments about agency (this is the reason Lewis gives). Second, it seems to undermine the concept of agency playing its intended role of (partly) grounding responsibility.

[37] Interestingly, Lewis' (1987, 187) discussion also involves a case akin to superiors' agency. See also the discussion in List and Menzies (2009, 497−98).

[38] I am indebted to a reviewer for suggesting this objection.

[39] Moreover, it leaves the over-generation worry unaddressed on the level of agency. Regardless of facts about a person's epistemic state, without the positive conditional, she would still be an agent of all downstream consequences (see Lewis, 1987, 184). This might weaken the explanatory contribution that agency

Second, the negative conditional raises a problem of too little agency (or: an under-generation problem) whereas the positive conditional is motivated by avoiding a problem of too much agency (or: an over-generation problem). Specifically, suppose that if Anne had *not* ordered Collin's death, someone else, say Ashley, would have. In fact, this should be so insofar as hierarchical groups have built-in redundancy. But then the second conditional seems false: Even if Anne had not ordered Collin's death, Ashley would have, and Collin would still die. This, I take it, is an instance of the well-known problem of redundant causation, or of a Frankfurt case with a counterfactual intervener.[41]

This problem arises when we think about Anne as a subordinate. Anne, although a superior to Bert and Ben, is—together with Ashley—a subordinate to someone else (because the group might have more than two levels). As far as this objection of redundant causation is concerned, Anne's position is like Bert's. Anne's order and Bert's shooting are similarly redundant. Collin would have died even if Bert had not shot (Ben would have). Likewise, an order to shoot Collin would have been given even if Anne had not given it (Ashley would have). Let's thus move on to subordinates' responsibility and return to the objection of redundant causation there. Here is just the very short answer: Following Peter van Inwagen (1983, 171–74), "Collin's death" should be distinguished into different events, such as Collin dying from *Bert's* shot, or him dying from *Anne's* order, etc. Hence, the backup commander Ashley, or the backup soldier Ben, do not make a difference to *these* results.

4.2 Subordinates' responsibility

To explain subordinates' responsibility—and to address this problem that all individuals in a hierarchical group are redundant but still make a difference—I need to introduce the second part of the account of agency in hierarchical groups: the individuation of outcomes.

Subordinates do not seem to make a difference because if they were not to follow a directive given to them, someone else would. Bert does not make a difference to

---

makes to responsibility (insofar as a relation that more narrowly circumscribes the grounds of an agent's responsibility gives a better explanation for why an agent is responsible when she is).

[40] The difference is that if Edward were to sharpen the knife, Giulia may or may *not* cut herself; but if Anne were to give the order, Collin will certainly be shot by someone.

[41] I am grateful to two anonymous reviewers for prompting me to discuss this problem.

Collin's death because, although the relevant positive conditional is true, the negative conditional is false. The positive conditional is true because Bert can make it the case that Collin dies reliably. In all relevantly similar situations, in which Bert intends to follow Anne's directive to shoot Collin, he goes ahead and shoots, and Collin dies. Yet, the negative conditional is false. Bert cannot prevent Collin's death. If Bert were to refuse Anne's directive, Bob or some other soldier in the team would go ahead and shoot Collin and Collin would die regardless. In this way, it is not up to Bert whether or not Collin dies. This leaves something to explain. On what grounds, then, is Bert responsible? I contend that there is something to which Bert makes a difference.

Although it is not up to Bert whether or not Collin dies, it is up to Bert whether or not *he* is the one who shoots Collin. If Bert were to refuse Anne's directive, then Collin would be shot by Ben, not by Bert. The result to which Bert makes a difference is thus *not* whether Collin dies, or whether Collin is shot, but whether Collin is shot *by Bert*.[42] For *this* outcome the negative conditional is true: If Bert were not to follow Anne's directive, then Collin would *not* be shot by Bert. Bert thus makes a difference to one outcome (that he, and not someone else, shoots Collin) but *not* another (that Collin is shot). In short, what Bert does, in virtue of which he might be responsible, is that *he* shoots Collin. Thus, subordinates' responsibility can be explained by distinguishing different outcomes.

We distinguish the outcome that Collin is shot from the outcomes that Collin is shot by Bert, and that Collin is shot by someone else. A similar line of reasoning is due to van Inwagen (1983, 171): That Caesar is murdered, that Caesar is stabbed, and that Caesar is poisoned are three different outcomes (or states of affairs).[43] That Caesar is poisoned and that he is stabbed each entail that he is murdered. The outcomes that he is poisoned and that he is stabbed are *more fine-grained* outcomes than that Caesar is murdered. Likewise, that Collin is shot by Bert, and that Collin is shot by someone else each entail that Collin is shot; and the first two outcomes (shot by Bert, shot by someone else) are more fine-grained than the outcome that Collin is shot. This idea can be made more precise if, following van Inwagen, we think of outcomes as sets of possible worlds. The more coarse-grained outcome that Collin is shot is represented by the set of all possible worlds in which Collin is shot, regardless of who shoots.[44] By contrast, the more fine-grained outcome of Collin being shot by Bert is represented by the subset of all possible worlds in which Collin is shot and in which, furthermore, the bullet that hit Collin came from the gun in Bert's hand.[45] More generally, an outcome is more fine-grained than another if it is its proper subset. Hence, it also makes sense to talk of outcomes being *nested*, with fine-grained outcomes being "nested in" coarse-grained outcomes.[46]

---

[42] This individuation might be clarified by seeing outcome descriptions as implicitly contrastive. Contrastivism is implied by the account here (List and Menzies 2009, 186–87; Sinnott-Armstrong, 2021), and well accepted in the literature (Schaffer, 2005; Sinnott-Armstrong, 2008; Menzies, 2008; see also Dretske, 1977). A similar non-contrastivist proposal of actions with "a layered structure" such that "distinct actions can have the same parts," is put forward by Ginet (1990, 50–51).

[43] More precisely: The three sentences "it is the case that Caesar…", what van Inwagen (1983, 171) calls "canonical names", express three distinct propositions—since "there are many different ways [the concrete particulars that make up our surroundings] could be arranged that would be sufficient for the obtaining of a given state of affairs."

[44] This set of possible worlds represents the outcome in that it holds fixed the occurrence of the outcome while varying everything else that happens "around it." I write "represent" to avoid suggesting that outcomes are identical to sets of possible worlds.

By thinking of outcomes as being represented by sets of possible worlds, the account becomes scalable. An outcome is fine-grained *compared to another* if it is its proper subset. Just as a hierarchical group has many levels (structured by the partial order of some authority relation), so outcomes have levels (structured by the partial order of the proper subset relation). The hierarchy of the group is thus mirrored in a hierarchy of outcomes. For each level in the hierarchy there is a corresponding level of grain of outcomes to which an individual at that level makes a difference.[47] A high-level individual in the hierarchy makes a difference to a corresponding coarse-grained outcome, a subordinate makes a difference to a fine-graining of this outcome. When hierarchical groups are conceived in this way—as a hierarchy of individuals, each of whom makes a difference to an outcome in a corresponding hierarchy of nested outcomes—then individual contributions become discernible.

This scalability now gives us the resources to consider Anne as a subordinate. The outcome to which Anne makes a difference is that Collin is shot *on her order*. If she were to give this order, Collin would be shot. If she were not to give this order, Collin might still be shot, but not on her order. If she were not to give this order, then Collin would *not* be shot on her order. That Collin is shot on her order (rather than on someone else's)—this is the outcome to which Anne makes a difference.

Anne does *not* make a difference to the outcome to which Bert makes a difference, namely, that Collin is shot by Bert. This is because, by assumption, if Anne were to give the order to shoot Collin, then *someone* would shoot Collin, but it is *not* the case that Bert follows Anne's order in all nearby possible worlds.[48] The outcome that Collin is shot *by Bert* is something that Anne can prevent (the negative conditional is true) but not something she can implement (the positive conditional is false).[49]

---

[45] This might prompt a circularity worry. The individuation of outcomes may seem to presuppose an understanding of agency although the outcomes are also used to define agency. There is no circularity here. The particular outcome that Collin is shot by Bert can be described without reference to agency. The addition "by Bert" can be spelled out without reference to intentions in purely behavioral terms. For example, the "by Bert" in this case can be spelled out as Bert's finger moving or as Collin's body being hit by a bullet from Bert's gun. As long as a non-agential analysis of this "by" addition is available, the particular outcome can be distinguished from the universal outcome without presupposing agency.

[46] These ideas are more rigorously expressed with the following three conditions. (1) *Outcomes*: An outcome is represented by the set of possible worlds in which it occurs. (2) *Distinctness*: If two outcomes are represented by different sets of possible worlds, then the two outcomes are distinct, i.e., not identical. (3) *Nestedness*: An outcome a is nested within an outcome b if and only if the set representing a is a strict subset of the set representing b.

[47] Paralleling the assumption that there are no overlapping competences in the hierarchy (i.e. each subordinate stands in the authority relation to exactly one superior), I make the following assumption requiring non-overlapping outcomes. Representing outcomes as sets, the intersection of any two outcomes is empty unless one outcome is a subset of the other.

[48] See footnote 29 above.

[49] Suppose, instead, that Bert shoots in *all* nearby possible worlds in which Anne gives the order. Anne then *would* make a difference to the outcome that Collin is shot by Bert. Such a case, that lacks redundancy (but keeps the characteristics of authority and functional dependence), would then be a case of *proxy*

One might object that this entails the implausible view that, because she does not make a difference to this outcome, Anne is not responsible for the outcome that Collin is shot by Bert. Suppose (a) that Anne *knows* that Bert is going to carry out her order, or only (b) that she knows that he is *likely* to carry out the order. Either way, it seems that, intuitively, when she gives the order to shoot Collin, she would be responsible for the outcome that Collin is shot *by Bert*—but the account here denies this.[50]

Before considering this objection, note that in one version, (a), the imagined case seems inconsistent. By assumption, Anne does *not* make a difference to the outcome that Collin is shot by Bert.[51] Yet, or so goes the case, Anne *knows* that Bert is going to be the one who shoots. This is inconsistent, insofar as knowledge is factive but Bert is not necessarily going to be the one who shoots. Given only that *someone* is going to shoot, but that this might not be Bert, Anne might give the order in the (mistaken) *belief* that Bert is going to shoot. Admittedly, Anne's responsibility might change between having a mistaken belief (that *Bert* will shoot) instead of a correct belief (that *someone* will shoot), but not in a way inconsistent with the account here.

To the other version of the objection, (b), I see two responses. First, the intuitive-seeming judgment that Anne is responsible for the outcome that Collin is shot by Bert might be denied. For example, internalists about responsibility contend that agents are responsible only for their attitudes or quality of their will (e.g. Scanlon, 2015, 96)—if they are responsible *for* anything at all. Zimmerman (2002, 568 emphasis in original) argues "whether there is something *for* which one is responsible is immaterial; all that matters, fundamentally, is whether one *is* responsible." Notably, internalists arrive at this view after reflecting on cases, just like ours, about control, luck, or "tracing" consequences back to the agent (e.g. Zimmerman, 2002; Khoury, 2012; 2018). Their view is consistent with my account. There is a difference between the reasons for and the objects of someone's responsibility. My account concerns only the grounds of someone's responsibility. These grounds explain why someone is responsible (they are reasons), which does not entail that someone is responsible *for* these grounds (they are not the objects).

Second, the response that I favor: We could accommodate this intuitive judgment (that Anne is responsible for the outcome that Collin is shot by Bert) to some degree by "redirecting" it without changes to the account. Per my account, Anne does not make a difference to Collin being shot by Bert, but she makes a difference to an outcome in which it is highly likely that Collin is shot by Bert.[52] What the intuitive judgment picks up on in saying that "Anne is responsible for the outcome that Collin is shot by Bert", that is, what makes this judgment look plausible, is in fact just that Anne is responsible for the outcome that it is highly likely that Collin is shot by

---

action: Anne is the agent of an action that Bert carries out in her stead as if he were her "tool" (cf. Feinberg 1970b, 222; Copp, 1979, 177; Ludwig, 2014; Himmelreich, 2018, 208).

[50] I am grateful for an anonymous reviewer for suggesting objections along these lines.

[51] By assumption *of the objection* that is. Otherwise, Anne would be an agent of the outcome that Collin is shot by Bert and thus can be responsible for this outcome. Similarly, superiors are regularly responsible for specifics when they make a difference to these specifics.

[52] Consider here the conditional probability of the outcome that Collin is shot by Bert given that Anne has given the order to shoot Collin.

Bert. That is, the objection's judgment rests on claims about likely consequences or about normative facts (e.g. the wrongness of Anne's action), not on a judgment about ontology and metaphysics that discriminates different entities for which Anne's is responsible. In this sense, the account here can accommodate a claim in the vicinity of the initial intuitive judgment of the objection.

To take stock: On the one hand, we have the two conditionals that define when individuals make a difference to an outcome. On the other hand, I distinguished a fine-grained outcome, that Collin is shot by Bert, and a more coarse-grained outcome that Collin is shot on Anne's orders. Together, this explains how both Anne and Bert are each an agent of some outcome that entails Collin's death. Bert makes a difference to the more fine-grained outcome, which thereby (partially) grounds his responsibility. Likewise, Anne, considered as a subordinate, makes a difference to the outcome that Collin is shot on *her* order (as opposed to the order of Ashley, Anne's backup) and this hence (partially) grounds Anne's responsibility.

The account may appear contrived with descriptions such as "the outcome that Collin is shot by Bert." But it captures a very natural idea: Hierarchical groups constrain and enable individuals at the same time. Bert can decide not to shoot, but whether Collin is shot is not up to him. Anne can choose not to give the order, but whether such an order is given is not up to her. Subordinates can choose their role in how a directive plays out—but play out it will. They act within a sphere of possibility that is circumscribed from above. When subordinates comply, what they bring about typically entails the objective of the directive they received, even if whether this objective is achieved is not up to them. These intuitive ideas are formalized by a simple set-theoretic construction of nested outcomes. The canonical names or descriptions of these outcomes might appear contrived or overwrought, but the larger ideas behind them should be rather clear and intuitive.[53]

4.3 Increasing responsibility

Michael Walzer discusses increasing responsibility in a way that resonates with these ideas. Walzer observes that "we regard soldiers under orders as men whose acts are not entirely their own and whose liability for what they do is somehow diminished" (1977, 309). Walzer suggests that, compared to their superiors, soldiers lack agency. To explain increasing responsibility, it seems natural to explore this further: How should we understand that "what [soldiers] do is somehow diminished" and that their acts are "not entirely their own"?

I hypothesize that when comparing superiors and subordinates, a change in responsibility can be explained, at least in part, by a respective change in agency. When we assume that moral responsibility is not a brute fact but that facts about responsibility are grounded in other facts, then there must be something about these grounding facts that explains increasing responsibility, that is, why responsibility increases with hierarchy: Moral responsibility is partially grounded in agency. If agency increased with hierarchy, then it could explain why responsibility increases with hierarchy. By discerning individual contributions and comparing them to another, we can see that agency increases with hierarchy and that thereby responsibility may increase with hierarchy.[54]

---

[53] I am grateful to a reviewer for encouraging me to highlight the intuitiveness of the account here.

Before attempting to explain increasing responsibility, the concept of "degrees of responsibility" should be cleared up (cf. Braham and Hees 2009, 325). First, a note on terminology. What I say about "degrees" can be expressed similarly in terms of "worse actions." More substantively, I suggest representing degrees of responsibility as a function of those facts that ground responsibility. Of those facts that ground responsibility, some concern agency, some concern what the agent believes or knows, and others concern what the agent intends—relating to the agential condition, the epistemic condition, and the agent's intentions or quality of will.[55] Degrees of responsibility reflect this metaphysical structure of responsibility. The function that represents degrees of responsibility should represent in its components these conditions of responsibility. An agent's degree of responsibility for an outcome is a function of the agent's contribution to this outcome, the agent's beliefs about the consequences, and the agent's intentions. But how can responsibility be represented as a graded concept permitting of degrees at all?

Responsibility can be graded whenever it is implicitly comparative.[56] Degrees of responsibility are constructed by comparing one agent's responsibility for a fine-grained outcome with another agent's responsibility for a coarse-grained outcome. First, degrees of agency are normed to 1, or "full agency," for some agent and an outcome to which this agent makes a difference. Call this the *index outcome* and *index agent* respectively. This index agent can be an idealized or hypothetical agent. The index agent makes a difference, or controls, whether or not the index outcome occurs. The index agent is furthermore supposed to have all relevant information about the situation (for simplicity, we set aside the intentional condition). As a matter of stipulation, the index agent—with full control and full information—has full responsibility for this outcome. With this idealization, the index agent's degree of responsibility with respect to this outcome is normed to 1.

Second, similarly to responsibility *simpliciter*, the degree to which one is responsible depends on the degree to which the grounding facts are in place. One is responsible for an outcome to the extent that one has agency over it, one foresaw it, and intended it. Whenever one of the components of degrees of responsibility decreases, relative to the index agent, an agent is responsible to a lesser degree. Focusing on agency only, I assume that degrees of responsibility are monotonically decreasing in agency.[57] Suppose another agent foresaw the same index outcome but lacked full agency over it. That is, this other agent had the same information as the index agent but made a difference only to an outcome nested in the index outcome. In this case she would *not* have full responsibility. A decrease in the control condition (agency only over a nested outcome), all other things being equal, leads to a decrease in the degree to which she is responsible compared to the index agent with full agency and full information. Any agent who has less than full agency and less than full information with respect to this outcome has a degree of responsibility less than 1. In technical terms, the function representing degrees of responsibility depends monotonically on its components of agency and information. Inversely, responsibility increases with hierarchy.

---

[54] This coheres with an approach in philosophy of law. "Loss of proximity to the act is compensated by an increasing degree of organizational control by the leadership positions in the apparatus" (Roxin, 2011, 200). See also Walzer (1977, 316).

[55] Moreover, responsibility might have a *deontic condition*, someone's responsibility might depend on (or be partly grounded in) the deontic status of an action, the degree of its wrongfulness, or its "moral significance." This would be another component in the function representing degrees of responsibility. But for present purposes, this deontic condition, like the intentional condition, has to be set aside.

[56] The law often proceeds in a similarly comparative fashion (Lepora & Goodin, 2013, 99).

[57] Strictly speaking, representing agency as degrees requires revising my earlier definition of agency as difference-making. The sufficient condition of difference-making would have to be understood as setting

These constructions allow only for very limited comparisons. Comparisons are possible only between a superior and subordinates.[58] Or, similarly, between the index agent and the agents who make a difference to outcomes nested within the index outcome. The constructions of degrees of responsibility exploit the idea that agents of fine-grained outcomes have in *one* sense less control than agents of more coarse-grained outcomes.[59] The agential component of this function is then defined only on the outcomes nested within this more coarse-grained outcome. The account does not define or speak to comparisons between agents whose outcomes are not nested.

Another limitation is that, besides agency and information, all other grounds of responsibility are set aside. How exactly information and the other grounds are measured and how all grounds interact as components to compose an overall measure of degrees of responsibility cannot be answered here. Other explanations of why responsibility increases with hierarchy are not ruled out. Thus, the phenomenon is only that responsibility *sometimes* increases with hierarchy; I do *not* claim that responsibility always increases. In fact, we can think of situations in which responsibility decreases with hierarchy. In some groups, superiors might be responsible to a lesser degree than their subordinates.[60] Hence, it is *not* the case that agents who make a difference to a nested outcome—such as subordinates—*always* have less responsibility than their superiors. Neither is it the case that superiors could generally reduce their responsibility by giving more specific orders (and thus make a difference to a more fine-grained outcome). And in some groups no one may have full responsibility. I discuss some of these and other cases before concluding.

We can now apply these technical constructions to illustrate increasing responsibility in hierarchical groups. Consider again the case of Anne and Bert who are both involved in bringing about Collin's death. The coarse-grained index outcome,

---

the agency measure to 1. The necessary condition would have to be dropped to allow for partial or shared agency with respect to a larger outcome.

[58] They do not need to be *direct* superiors and subordinates, but there can be intermediate superiors or subordinates respectively (i.e., the two agents need to be connected in the authority relation that defines the hierarchy).

[59] The sense here is a relative power to prevent outcomes. There is a different sense of agency understood not as the power to prevent but rather as the power to fine-tune some result. On this latter sense of fine-tuning, subordinates generally have a higher degree of agency than their superiors.

[60] Suppose a mafia captain tells her associates to "get rid of Collin," without specifying how exactly they ought to do so. In fact, the associates could just tell Collin to leave Chicago forever, but they could also poison Collin, shoot him, or even torture him to death. Given their range of options, when opting for torture, the associates would be more responsible than their captain.

for comparing Anne and Bert, is Collin's death due to Anne's orders. Anne, as the index agent, counts as having "full agency" over this outcome in this comparison. Lesser degrees of agency are assigned to the agents of the more fine-grained, nested outcomes. Specifically, Bert, the agent of the nested outcome (that Collin is shot by Bert), has only partial agency with respect to the index outcome (that Collin's is shot on Anne's orders).[61] Hence, Anne's responsibility is greater than Bert.

These two ideas—the idea of degrees of agency (as comparing agency between two outcomes, whereas one is nested within the other) and that of degrees of responsibility (defined as a function that is monotonically increasing in agency)—together explain how responsibility increases with hierarchy.[62] It follows from these constructions that when other conditions of responsibility are set aside, then the degree of responsibility increases as the degree of agency increases. And since the degree of agency increases with hierarchy, responsibility may increase with hierarchy.

The construction on how responsibility increases with agency can moreover be extended to a case like the following.[63] Assume, again, our model case of Anne and her soldiers; but suppose now that two types of guns are available to Anne's soldiers, "normal" guns and "brutal" guns. Brutal guns inflict greater pain and a slower death. Suppose now that another of Anne's subordinates, Bill, can reliably get Bert to pick a brutal gun rather than a normal gun.[64] Say, Bill is always in the armory at the same time and chooses to stand in the way of Bert getting to the normal guns. In other words, Bill makes a difference to whether Collin is killed by Bert *with a brutal gun* (as opposed to a normal gun). Bert knows that there are two types of guns and what they are and chooses to not go around Bill to get to a normal gun. I take it that the pre-theoretic judgment is that Bill is *not* responsible; and if he is at all responsible, then he is less responsible than Bert (after all, it is Bert who shoots Collin).

The present account can make good sense of this case. First, the account can explain why Bill might be responsible (if he makes a difference to a relevant outcome) but the account does not entail that Bill *is* responsible. Agency is only a partial ground but not a sufficient condition, let alone a full explanation, for someone's responsibility. Second, the account entails only the claim that, if Bill is responsible at all (assuming he and Bert have the same information), then Bill is responsible to a lesser degree than Bert. The account identifies in what sense Bill has less agency than Bert—he makes a difference to a more fine-grained outcome (which gun Bert uses)—which hence grounds his lesser responsibility. The account thus extends to some further and interesting cases.

---

[61] See footnote 59.

[62] The claim is *not* that this is the only explanation. Another explanation for increasing responsibility could be that the actions "get worse" or more significant as we go up the hierarchy (as suggested by one reviewer). However, whether this ("responsibility increases with significance") is an independent ground or one that partly depends on ideas such as that of nested actions (or that "responsibility increases with agency"), cannot be settled here. And it need not, since the two explanations seem compatible.

[63] I am indebted to an anonymous reviewer for this observation and for suggesting a case along these lines.

[64] This assumption might be a stretch. Arguably, there are nearby possible worlds where Bert is about to shoot Collin, but Bill fails to get Bert to pick a brutal gun. However, without this assumption, this case would be easy to explain: Without this assumption, Bill doesn't make a difference (to an outcome that entails Collin's death). The aim is to explain how Bill is not responsible for Collin's death despite making a difference. This assumption ensures that Bill *makes* a difference to an outcome that entails Collin's death.

5 Observations, limitations, and extensions

So far, I have concentrated on the case of Anne, Bert, and other soldiers in Anne's team. I have taken this case as a model under the assumption that a theory developed for this model can be extended to more complex cases. More complex cases will involve groups with more than two levels, any number of subordinates for a given superior, and the possibility to resist directives. More complex cases will also have groups that combine vertical and horizontal elements. A corporation, for example, is headed by a board of directors instead of having one individual at the top. Such a board is a horizontal group. Even a group agent could occupy any node in the hierarchy tree. In contrast, the group in our model case is *strictly hierarchical*, that is, the authority relation that represents the group's structure relates only individuals. Only individuals occupy the nodes in the hierarchy tree. One individual is at the top and she is connected to all other group members individually via a path in the authority relation.

The simplifications of the model case should not affect the generality of the account presented here. They do not in principle stand in the way of extending this account to more complex cases. In considering Anne in her role as a subordinate, I already discussed how the proposal can scale to more levels. Before closing, to both to clarify the account and to connect it to related topics, I want to illustrate some other ways in which this theory of agency in hierarchical groups may extend to more complex cases and related discussions in the literature. I discuss three issues: ignorance, individual responsibility, and collective responsibility.

5.1 Ignorance and plausible deniability

Responsibility does not always increase with hierarchy. This is because responsibility is only partially grounded in agency and hence agency is by itself not sufficient for moral responsibility. All conditions of responsibility matter. Given how action, knowledge, and intent come together, one may argue, that responsibility sometimes *decreases* with hierarchy. Some groups shape their internal dissemination of information strategically to absolve superiors of their moral responsibility. Subordinates may deliberately withhold information from superiors knowing that if their superiors were to have this information, the superiors would be responsible. So, subordinates may arrange things hoping that their superiors, despite their agency but because of their ignorance, might not be responsible, or cannot be held responsible.[65] This is sometimes called *plausible deniability*. The idea is, of course, that pleading ignorance avoids responsibility. It is doubtful that this is the case (cf. Kutz, 2000, 157).

---

[65] That subordinates intend to shield their superiors suggests an important relation between this way of aiming for plausible deniability and complicity. Subordinates and superiors might be complicit in this scheme.

In other cases, responsibility may decrease with hierarchy because disseminating relevant information to the top is not feasible. Consider a military group. Soldiers in the field often know better than their commanders what is going on. Yet, given the conditions of battle, this information cannot travel upwards. In result, the group is control-heavy at the top but information-heavy at the bottom. In technical terms, the degree of information is inversely related to the degree of agency. Although agency increases with hierarchy, information may decrease. Crucially, in contrast to plausible deniability, there is no intention here to avoid sending information up the hierarchy. And hence, at least in some cases, finding anyone who is fully morally responsible in such a group might be hard. Moral responsibility requires agency *and* information to come together in some sense. Although the individual contributions to the collective outcome may be discernible, if agency and information do not sufficiently line up, then agents might escape responsibility, or so one may argue.

These suggestions that responsibility might decrease with hierarchy should not hide the fact that information, hierarchies, and responsibility interact in complex ways. Whether and when ignorance diminishes responsibility is not relevant for the account presented here. The account can accommodate judgments either way. Where information diminishes with hierarchy, thinking of degrees of responsibility as this account suggests—as a function that has the grounds of responsibility as components—can be useful and extended to investigations into the epistemic condition of responsibility in hierarchical groups. At any rate, even if responsibility decreases with ignorance and hierarchy, ignorant superiors will escape responsibility of only one kind (the one discussed here). They might still need to take responsibility, be answerable, or perhaps can still be held responsible (depending on the theory of responsibility).

5.2 Individual responsibility and command responsibility

My account suggests that in hierarchical groups individuals are responsible. In a strictly hierarchical group, we can discern the individual contributions to some overall outcome. Even if a subordinate cannot make a difference to a coarse-grained outcome, she, via a fine-grained outcome, still makes a difference as to how this outcome comes about. Subordinates are agents of nested outcomes. Given that the individuals have the relevant information, they can be individually responsible in virtue of their individual contributions. In a strictly hierarchical group, for each individual there is an outcome to which that individual makes a difference. Moreover, in a strictly hierarchical group, there are only individuals making a difference to outcomes. I do not deny that also hierarchical groups might have corporate responsibility.[66] Rather, this corporate responsibility need not be a "remainder of responsibility" (Dempsey, 2013, 343), and it does not result from a "deficit in the accounting books" (Pettit, 2007, 194).[67]

---

[66] Specifically, my view is consistent with the "two-level view" that distinguishes between questions of individual and collective responsibility. See Isaacs (2011).

[67] Similarly, Kutz (2000, 113) writes that there is often a "disparity between collective harm and individual effect… If no individual makes a difference, then no individual is accountable for these collective harms."

An important application for an account of individual responsibility in hierarchical groups is *command responsibility*. "Command responsibility" names a family of doctrines in criminal law based on which superiors are held responsible for actions performed on their command. The development of such doctrines goes back to the conviction of the Japanese general Tomoyuki Yamashita for the Manila massacre in 1946 (Walzer, 1977, Chap. 19; Smidt, 2000; Danner and Martinez 2005). Subsequent to the Second World War, several different doctrines of command responsibility have been developed. German criminal law, for example, developed the doctrine of indirect perpetration (Roxin, 1963, 2011; Weigend, 2011).

This account of agency in hierarchical groups relates to doctrines of command responsibility in two ways. First, by explaining superiors' responsibility, this account of agency in hierarchal groups vindicates an idea close to the basic tenet of command responsibility.[68] Commanders are related to the outcome for which they are responsible in a way that grounds their responsibility for these outcomes. Second, the account also highlights the limits of commanders' responsibility. Although commanders may make the difference with respect to a coarse-grained outcome, they still lack the ability to fine-tune (which lies with their soldiers). In other words, commanders lack the implementation aspect of agency. Moreover, commanders may also be lacking (the possibility of) information about the consequences of their order. This kind of ignorance may also affect their responsibility. When responsibility is seen as a function of agency and information, then commanders may be long on agency (at least as prevention, if not as implementation) but short on information. Assuming the validity of plausible deniability (that ignorance undermines responsibility), this may pose a challenge to founding the legal doctrine of command responsibility on moral responsibility.

Finally, reflecting on doctrines of command responsibility brings out an important limitation of the account of hierarchical groups presented here. The present account is concerned only with agency *within* a hierarchical structure, not with agency *over* this structure. Superiors are not seen as being the ones who set up the organization and design the procedures by which it operates. The power structure is taken as given. Yet power over an organization is a crucial aspect of command responsibility. This is an aspect on which my account is silent and that needs to be developed another day.

5.3 Mutiny as a collective action

It is now time to relax one assumption I made earlier about the model case. Let us assume a mutiny is possible. In contrast to my earlier assumption that it is certain that Collin will be shot by someone when Anne gives the order, let us say that Collin being shot is only overwhelmingly likely. There is also the small possibility that no one will shoot. Collin might survive. But Collin would survive only when each soldier in the team refused to follow Anne's order. Note how the situation resembles that of a collective action: Suppose five of us need to rescue a child drowning at the

---

Pace Kutz, on my account there are groups where individuals do make a difference.

[68] Only "an idea close to" because it does not vindicate the idea, which could taken to be the basic tenet of command responsibility, that individuals are responsible for a *collective* action.

beach and all five of us are needed to get it done. We need to get our act together. Individually none of us can ensure that the child is rescued. If we fail to get our act together, then we might be individually responsible for not taking the appropriate steps towards a collective action, and collectively we might be responsible for not rescuing the child (Collins, 2012). We can now see how a similar collective action problem befalls the team of soldiers.

What are soldiers in our model case responsible for? A mutiny is a collective action. Because one person is enough to frustrate the effort to resist Anne's order, each soldier would have to resist her order. Hence, individually each soldier might be responsible for not taking appropriate steps towards a collective resistance, and *collectively* the soldiers might be responsible for the omission of not resisting Anne's order together (this depends on a theory of responsibility in horizontal groups).

Throughout this paper I have assumed a kind of local determinism according to which Collin is shot when Anne directs he be shot, and according to which a bullet hits Collin fatally when Bert pulls the trigger. But in reality, luck is involved. In reality, actions are risky. Technically, we can represent risk by thinking of probabilistic outcomes. The outcome that Collin is shot, to which Anne makes a difference, includes not only worlds in which Collin is shot, but also some few worlds in which there is a mutiny. Likewise, the outcome that Collin is shot by Bert includes also some few worlds in which Bert misses his target. Of course, more would have to be said on how my account extends to such probabilistic outcomes.

6 Conclusion

Many groups—banks, ministries, corporations, or military organizations—are hierarchical groups structured by authority. Who is responsible for the conduct of such a group?

Responsibility has many grounds. I concentrated on the partial ground of control, causation, or agency. Agency identifies the things in virtue of which an individual may be responsible. Agency explains why we are responsible for our own actions but usually not for those of others. A theory of agency in hierarchical groups is an important building block of a wider theory of responsibility in hierarchical groups. The central challenge to theories of responsibility in collective contexts is to discern who brought about what and who, accordingly, can be held responsible for what.

In this paper, I have done three things. First, I contrasted hierarchical groups to flat groups by identifying three characteristics of hierarchical groups: asymmetry, control, and redundancy. Second, I described three phenomena that any account of hierarchical groups needs to explain: superiors' responsibility, subordinates' responsibility, and increasing responsibility. Third, I put forward a theory of agency in hierarchical groups that consists of two parts. The first part is an agency relation that relates individuals to outcomes. Individuals stand in this relation when something is up to them, when they make a difference: when they can both robustly bring about and prevent an outcome. In spelling this out, I have relied on an existing account of difference-making, due to List and Menzies (2009). The second part is a proposal about the individuation of outcomes. I suggested distinguishing outcomes contrastively. Outcomes can be nested in other outcomes and a superior can make a difference to a coarse-grained outcome while subordinates make a difference to nested, fine-grained outcomes. These two parts of the theory explain the first two phenomena, superiors' responsibility and subordinates' responsibility. The third phenomenon, increasing responsibility, can be explained by constructing a comparative measure of degrees of responsibility that is increasing in agency. Responsibility may increase with agency, setting other things aside, because agency increases with hierarchy.

More work is needed to couch this view within existing theories. Foundations need to be explored to theories below. How does this theory of agency relate to the metaphysics of causation and, more broadly, to the philosophy of action? Lines must be drawn to theories above. Would such a theory of responsibility in hierarchical groups buttress or undermine existing doctrines in criminal law? Finally, connections need to be established left and right. How do the intentional and the epistemic condition of responsibility interact with the agency condition? The unification that I attempted for responsibility in hierarchical groups is only one colorful piece in the large mosaic of theories.

Acknowledgements This paper sees the light of day a little over 10 years since it was first presented at the London School of Economics and at the University of London. It has benefitted from discussions at the Graduate Reading Retreat of the Stockholm Centre for the Ethics of War and Peace, at the Humboldt Universität Berlin Philosophy Colloquium, Neil Roughley's colloquium in Duisburg-Essen, and at Collective Intentionality X. The many people who commented on drafts or helped in conversations include Saba Bazargan-Forward, Gert Keil, Gabriel Wollner, Walter Sinnott-Armstrong, Matthew Adler, and Deborah A. DeMott, Christian List, Susanne Burri, Helen Frowe, Ying Shi, Jesse Saloom, Sebastian Köhler, and Seth Lazar, as well as three anonymous referees for this journal (one of whom in particular saved me from many errors large and small).

Declarations

Consent for publication The paper has not been published previously in any form and it is not currently under consideration for publication elsewhere.

References

Bazargan-Forward, S. (2017). Accountability and intervening agency: An asymmetry between upstream and downstream actors. *Utilitas*, *29*(1), 110–124. https://doi.org/10.1017/S0953820816000224

Bazargan-Forward, S. (2022). *Authority, cooperation, and accountability*. Oxford University Press.

Braham, M., & Martin van, H. (2009). Degrees of causation. *Erkenntnis*, *71*(3), 323–344. https://doi.org/10.1007/s10670-009-9184-8

Braham, M., & Martin van, H. (2012). An anatomy of moral responsibility. *Mind*, *121*(483), 601–634. https://doi.org/10.1093/mind/fzs081

Bratman, M. E. (2006). What is the accordion effect? *The Journal of Ethics*, *10*(1/2), 5–19.

Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford University Press.

Buchanan, A. (1996). Toward a theory of the ethics of bureaucratic organizations. *Business Ethics Quarterly*, *6*(4), 419–440. https://doi.org/10.2307/3857497

Collins, S. (2012). Collectives' duties and collectivization duties. *Australasian Journal of Philosophy*, *91*(2), 231–248. https://doi.org/10.1080/00048402.2012.717533

Copp, D. (1979). Collective actions and secondary actions. *American Philosophical Quarterly*, *16*(3), 177–186. https://doi.org/10.2307/20009757

Crawford, N. C. (2007). Individual and collective moral responsibility for systemic military atrocity. *Journal of Political Philosophy*, *15*(2), 187–212. https://doi.org/10.1111/j.1467-9760.2007.00278.x

Danner, A., Marston, & Martinez, J. S. (2005). Guilty associations: Joint criminal enterprise, command responsibility, and the development of International Criminal Law. *California Law Review*, *75*, 169.

DeFalco, R. C. (2013). Contextualizing Actus Reus under Article 25(3)(d) of the ICC Statute thresholds of contribution. *Journal of International Criminal Justice*, *11*(4), 715–735. https://doi.org/10.1093/jicj/mqt045

DeMott, D. A. (2014). The fiduciary character of agency and the interpretation of instructions. In A. S. Gold & P. B. Miller (Eds.), *Philosophical foundations of fiduciary law* (1st edn.). Oxford University Press.

Dempsey, J. (2013). Corporations and non-agential moral responsibility. *Journal of Applied Philosophy*, *30*(4), 334–350. https://doi.org/10.1111/japp.12029

Dretske, F. I. (1977). Referring to events. *Midwest Studies in Philosophy*, *2*(1), 90–99. https://doi.org/10.1111/j.1475-4975.1977.tb00030.x

Eldar, S. (2013). Indirect co-perpetration. *Criminal Law and Philosophy*, 1–13. https://doi.org/10.1007/s11572-013-9224-y

Estlund, D. (2007). On following orders in an unjust War. *Journal of Political Philosophy*, *15*(2), 213–234. https://doi.org/10.1111/j.1467-9760.2007.00277.x

Feinberg, J. (1970a). Action and responsibility. *Doing and deserving: Essays in the theory of responsibility* (pp. 119–151). Princeton University Press.

Feinberg, J. (1970b). Collective responsibility. In *Doing and deserving: Essays in the theory of responsibility* (pp. 222–251). Princeton University Press.

French, P. A. (1979). The corporation as a moral person. *American Philosophical Quarterly*, *16*(3), 207–215.

Ginet, C. (1990). *On action*. Cambridge University Press.

Hall, N. (2004). Two concepts of causation. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 225–276). The MIT Press.

Heath, J. (2020). *The machinery of government: Public administration and the liberal state*. Oxford University Press.

Hertle, H. H., & Nooke, M. (2009). *Die Todesopfer an Der Berliner Mauer 1961–1989: Ein biographisches Handbuch*. Ch. Links Verlag.

Himmelreich, J. (2018). Agency and embodiment: Groups, human–machine interactions, and virtual realities. *Ratio*, *31*(2), 197–213. https://doi.org/10.1111/rati.12158

Himmelreich, J. (2023). What killed your plant? Profligate omissions and weak centering. *Erkenntnis*, *88*(April), 1683–1703. https://doi.org/10.1007/s10670-021-00422-9

Inwagen, P. (1983). *An essay on Free Will*. Clarendon Press.

Isaacs, T. (2011). *Moral responsibility in collective contexts*. Oxford University Press.

Khoury, A. C. (2012). Responsibility, tracing, and consequences. *Canadian Journal of Philosophy*, *42*(3–4), 187–207. https://doi.org/10.1080/00455091.2012.10716774

Khoury, Andrew C. 2012. (2018). The objects of moral responsibility. *Philosophical Studies* 175 (6): 1357–1381. https://doi.org/10.1007/s11098-017-0914-5

Kutz, C. (2000). *Complicity: Ethics and law for a collective age*. Cambridge University Press.

Lawson, B. (2011). Individual complicity in collective wrongdoing. *Ethical Theory and Moral Practice*, *16*(2), 227–243. https://doi.org/10.1007/s10677-011-9332-5

Lepora, C., & Goodin, R. E. (2013). *On complicity and compromise*. Oxford University Press.

Lewis, D. (1973a). *Counterfactuals*. Wiley-Blackwell.

Lewis, D. (1973b). Causation. *The Journal of Philosophy*, *70*(17), 556–567. https://doi.org/10.2307/2025310

Lewis, D. (1987). Postscripts to ‘causation’. In *Philosophical papers: Volume II* (pp. 172–213). Oxford University Press.

Lipsky, M. (1980). *Street-level bureaucracy: The dilemmas of the individual in public service*. Russell Sage Foundation.

List, C., & Menzies, P. (2009). Non-reductive physicalism and the limits of the exclusion principle. *The Journal of Philosophy*, *106*(9), 475–502.

List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press.

Ludwig, K. (2014). Proxy agency in collective action. *Noûs*, *48*(1), 75–105. https://doi.org/10.1111/nous.12013

Manacorda, S., & Meloni, C. (2011). Indirect perpetration versus Joint Criminal Enterprise Concurring approaches in the practice of International Criminal Law? *Journal of International Criminal Justice*, *9*(1), 159–178. https://doi.org/10.1093/jicj/mqq074

May, L. (1987). *The morality of groups*. University of Notre Dame Press.

May, L. (1992). *Sharing responsibility*. University of Chicago Press.

May, L. (2005). *Crimes against humanity: A normative account*. Cambridge University Press.

Menzies, P. (2008). The exclusion problem, the determination relation, and contrastive causation. In J. Hohwy & J. Kallestrup (Eds.), *Being reduced: New essays on reduction, explanation, and causation* (pp. 196–217). Oxford University Press.

Menzies, P. (2011). The role of counterfactual dependence in causal judgements. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.), *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 186–207). Oxford University Press.

Pettit, P. (2007). Responsibility incorporated. *Ethics*, *117*(2), 171–201.

Prosecutor v. Katanga and Chui (2008). International Criminal Court.

Rosen, G. (2010). Causation, counterfactual dependence and culpability: Moral philosophy in Michael Moore's causation and responsibility. *Rutgers Law Journal*, *42*, 405.

Rosen, G. (2015). The alethic conception of moral responsibility. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The nature of moral responsibility: New essays* (pp. 65–88). Oxford University Press.

Roth, A. S. (2014). Indispensability, the discursive dilemma, and groups with minds of their own. In S. R. Chant, F. Hindriks, & G. Preyer (Eds.), *From individual to collective intentionality* (pp. 136–162). Oxford University Press.

Roxin, C. (1963). *Täterschaft Und Tatherrschaft*. Cram, de Gruyter & Co.

Roxin, C. (2011). Crimes as part of organized power structures. *Journal of International Criminal Justice*, *9*(1), 191–205. https://doi.org/10.1093/jicj/mqq083

Sartorio, C. (2004). How to be responsible for something without causing it. *Philosophical Perspectives*, *18*(1), 315–336. https://doi.org/10.1111/j.1520-8583.2004.00031.x

Scanlon, T. (2015). Forms and conditions of responsibility. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The nature of moral responsibility: New essays*. Oxford University Press.

Schaffer, J. (2005). Contrastive causation. *The Philosophical Review*, 327–358.

Searle, J. R. (2010). *Making the social world*. Oxford University Press.

Shapiro, S. (2014). Massively shared agency. In M. Vargas, & G. Yaffe (Eds.), *Rational and social agency: The philosophy of Michael Bratman* (p. 257). Oxford University Press.

Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics*, *121*(3), 602–632.

Shoemaker, D. (2015). *Responsibility from the margins*. Oxford University Press.

Sinnott-Armstrong, W. (2008). A contrastivist manifesto. *Social Epistemology*, *22*(3), 257–270. https://doi.org/10.1080/02691720802546120

Sinnott-Armstrong, W. (2021). Contrastive mental causation. *Synthese*, *198*(3), 861–883. https://doi.org/10.1007/s11229-019-02506-0

Smidt, M. L. (2000). Yamashita, Medina, and beyond: Command responsibility in contemporary military operations. *Military Law Review*, *164*, 155–234.

Smith, A. M. (2006). Control, responsibility, and moral assessment. *Philosophical Studies*, *138*(3), 367–392. https://doi.org/10.1007/s11098-006-9048-x

Sunstein, C. R., & Edna, U. M. (1999). Second-order decisions. *Ethics*, *110*(1), 5–31. https://doi.org/10.1086/233202

Szigeti, A. (2014). Collective responsibility and group-control. In J. Zahle, & F. Collin (Eds.), *Rethinking the individualism-holism debate*. Synthese Library 372. (pp. 97–116). Springer International Publishing.

The Attorney General v. Eichmann (1961). Jerusalem District Court.

Vest, H. (2014). Problems of participation — Unitarian, differentiated approach, or something else? *Journal of International Criminal Justice*, *12*(2), 295–309. https://doi.org/10.1093/jicj/mqu021

Walzer, M. (1977). *Just and unjust wars: A moral argument with historical illustrations*. Basic Books.

Walzer, M. (2004). Two kinds of military responsibility. In *Arguing about war*. Yale University Press.

Wasserstrom, R. A. (1980). Conduct and responsibility in war. *Philosophy and Social issues: Five studies*. University of Notre Dame Press.

Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, *24*(2), 227–248. https://doi.org/10.5840/philtopics199624222

Weigend, T. (2011). Perpetration through an organization the unexpected career of a German legal concept. *Journal of International Criminal Justice*, *9*(1), 91–111. https://doi.org/10.1093/jicj/mqq077

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review*, *115*(1), 1–50.

Zimmerman, M. J. (2002). Taking luck seriously. *The Journal of Philosophy*, *99*(11), 553. https://doi.org/10.2307/3655750

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.