ORIGINAL PAPER

Mapping the Contours of Blame: An Account of the Moral Boundaries of Organizations

Rita Mota[1] · Alan D. Morrison[2]

Abstract

This paper presents an account of the moral boundaries of organizations. We define an organization's moral boundary to encompass all of the actions for which it could be held morally responsible. Our theory requires us to view organizations as subjects that act in the world, rather than as objects that are used as tools; that is, it requires us to focus on corporate moral agency. We present a process model for determining whether a given action lies within an organization's moral boundary, and we discuss how an organization's moral boundary can be created, destroyed, or modified as a result of deliberate choices by human and organizational actors. Our article contributes to the literature by conceptualizing the distinction between organizations as subjects and organizations as objects, and so clarifying the distinction between legal and moral boundaries; by recentering the discussion of boundaries on organizational actions rather than on contingent institutional features; and by adding nuance to the assignment of moral responsibility in complex organizational networks and in situations where one corporate moral agent depends upon another for its existence.

Keywords Organizational boundaries · Corporate moral agency · Corporate moral responsibility · Nested moral agency

Introduction

Blame for unethical behavior by a local subsidiary of a multinational enterprise is often ascribed to the parent firm. For example, Royal Dutch Shell (RDS) was blamed for human rights abuses and environmental damage perpetrated by its Nigerian subsidiary (see Reed, 2021). Situations like this raise difficult questions: from a legal point of view, Shell Petroleum Development Company (SPDC) of Nigeria is distinct from RDS so that, for example, creditors of SPDC have no claim upon RDS in the event of SPDC's bankruptcy.

And yet many people feel that it is natural to blame RDS for actions that are legally attributable to SPDC. This paper presents a framework within which it is possible to study the contours of corporate moral responsibility in cases like this one. That is, we present a specifically *moral* theory of organizational boundaries that allows us to identify the proper locus of moral responsibility in complex situations of this type.

Substantial literatures in economics, law, and management study the place where the organization stops and the rest of the social world starts; while those literatures advance our understanding of the corporate form, none of them is intended to address explicitly moral questions. Rather, they define organizational boundaries in order to study instrumental goals relating to efficiency, policy, and control.[1] The economic literature on organizational boundaries can be traced back to Coase's (1937) seminal analysis. Coase and the literature that develops his ideas are concerned with the relative merits of resource allocation by conscious control and decentralized prices, and claim that organizations stop at the point where

✉ Rita Mota
rita.mota@esade.edu

Alan D. Morrison
alan.morrison@sbs.ox.ac.uk

[1] ESADE Business School, Ramon Llull University, Av. Torreblanca, 59, 08172 Sant Cugat, Barcelona, Spain

[2] Saïd Business School and Merton College, University of Oxford, Park End Street, Oxford OX1 1HP, UK

---

[1] The literature on this subject typically refers to the "boundaries of the firm." We are concerned with all types of organizations and, hence, we refer throughout this article to "organizational boundaries."

managerial direction is displaced by market clearing. The legal literature understands organizations in terms of the assets they own and the contracts they write; it identifies several organizational boundaries, according to the functional use to which they will be put. For example, the regulation of going-concern companies and the winding up of failed enterprises each have a different goal and, hence, require a different understanding of organizational boundaries (Armour et al., 2017; Hansmann & Kraakman, 2000; Hansmann et al., 2005). And, in order to understand the organization and sustainability of cooperative activity, management scholars conceptualize organizational boundaries in terms of the locus of power relationships, the loyalties of employees, and the ways in which competition and cooperation occur (e.g., Araujo et al., 2003; Barney, 1991; Kast & Rosenzweig, 1972; Sabel, 1993; Womack et al., 2007).

The various literatures on organizational boundaries have generated important insights into the most efficient way to structure productive activity and its regulation, but they are not concerned with fundamental questions about the organization's moral responsibilities. We therefore pose the following questions. First, how should organizations be conceived in order that their moral boundaries can be meaningfully conceptualized? Second, how does one decide whether a given action falls within an organization's moral boundary? Third, how can a given organization change its moral boundary?

Our analysis yields the following conclusions. First, we argue that, in order to define an organization's moral boundary, it is necessary that the organization be understood as a subject that acts in the world, rather than as an object that is used to accomplish other subjects' goals. That subject must be the sort of thing that can reasonably be held responsible for its actions: that is, it must be a *moral agent*. Only a moral agent can have a moral boundary; we conceptualize that boundary as encompassing the actions that can be morally attributable to the agent. There is no a priori reason to believe that an organization's moral boundary coincides with the Coasean efficiency-oriented boundary or with the legal boundary used when organizations are wound up or sued, because the unit of analysis is radically different. In order to address our second question, we present a process model that explains how a given organizational action is attributed to one or more organizations, or to one or more human agents. Moral attribution to an organization is appropriate only if it possesses moral agency. If it does not, then we argue that moral responsibility must be assigned either upwards, to a controlling organization, or downwards, to individuals. Third, an organization's moral boundary can be created, destroyed, or modified. Creation and destruction are associated with acquisition and loss of moral agency; modification occurs

when the organization changes the actions that it is capable of performing.

Our approach frees us from some of the tacit assumptions underpinning prior accounts of organizational boundaries: we can identify moral boundaries even for organizations that do not attempt to maximize efficiency, irrespective of the goals of bankruptcy policy, and without regard to the way that going-concerns are regulated. But we free ourselves from these assumptions at a cost: our account succeeds only to the extent that organizations *are* moral agents; in this article, we assume that they can be.

The article proceeds as follows. The following section outlines the existing literatures on organizational boundaries and explains that, because those literatures view organizations as objects, they cannot answer questions about moral responsibility. We then discuss what it means to view an organization as a subject, and we survey different approaches to corporate moral agency. We then define the moral boundary of an organization, and we present a process model for determining whether an action sits within that boundary. We discuss the ways in which an organization's moral boundary could be created, destroyed, or modified, and we conclude by highlighting our contributions and identifying possible future research directions.

The Organization as an Object

This section outlines existing understandings of organizational boundaries. Every relevant account of where organizations end and the rest of the social world begins views the organization as an object in the world and explains its boundary in terms of the uses to which people want to put it. We review the dominant strands of literature, and we argue that boundaries drawn from an object-based perspective are a poor foundation for arguments about corporate moral responsibility.

The most influential accounts of organizational boundaries start from the position that organizations are devices for allocating resources efficiently and make the positivistic claim that their boundaries are drawn in a way that maximizes economic efficiency. The seminal work in this literature is due to Coase (1937), who sees organizations as objects that people use to reduce the costs of economic cooperation. According to Coase, organizational boundaries separate resources that are consciously allocated by managers from those that are allocated in impersonal arm's-length markets; he famously invokes Robertson's (1923) description of organizations as "islands of conscious power in this ocean of unconscious cooperation like lumps of butter coagulating in a pail of buttermilk." Coase argues that market forces determine organizational boundaries in a way that minimizes transactions costs. This analysis informs

a swathe of theoretical and empirical work. For example, Williamson (1975, 1985) characterizes the relative costs of transactions within and without markets and claims that they can be used to explain the choice between organizational structures, the scope of the organization, and the institutions that support market activity; and Grossman and Hart (1986) and Hart and Moore (1990) argue that the topology of organizational boundaries minimizes the deadweight costs of opportunistic hold-up behavior when incomplete contracts are renegotiated.

Some important normative contributions to the management literature are inspired by Coasean positivistic analysis. For example, the Resource-Based View of the Firm conceptualizes the organization as a collection of assets and asks how its boundaries should be drawn in order to achieve sustained competitive advantage (Barney, 1991; Barney et al., 2001; Mahoney & Rajendran Pandian, 1992; Wernerfelt, 1984). Relatedly, the literature on lean production asks how organizational boundaries ought to be positioned in pursuit of efficient supply chain management (e.g., Womack et al., 2007).

Legal scholars use Coasean transaction costs analysis in order to identify the most cost-effective way to pursue policy objectives. When the organization is a going concern (that is, when it will continue to operate for the foreseeable future), it is treated as a device for allocating resources, and the law is designed to reduce the contracting costs relating to information asymmetries and moral hazard. In this case, in line with Coase, the organization's boundary is assumed to lie where managerial and investor control give way to market forces (see, e.g., Armour et al., 2017). In contrast, a bankrupt corporation in winding-up procedures is no longer operating and its boundary therefore does not reflect control rights that are exercised for the long term; instead, the boundaries are drawn in order that property rights over its assets can be enforced as efficiently as possible in light of the principles of fairness embodied in the associated legal regime (Hansmann & Kraakman, 2000; Hansmann et al., 2005).

The literature that we have surveyed views organizations as objects that are used in pursuit of a particular goal, and evaluates an organization and its boundary according to the contribution that they make to that goal. This type of evaluation is analogous to judgments of a hammer in terms of its efficacy as a device for driving nails into wood. For example, Special Purpose Entities (SPEs) are created by drawing a legal boundary around a collection of assets that is controlled by another legal entity. The original goal of legally separating the assets held by an SPE from its parent was to increase the efficiency of property rights allocation by preventing the parent's creditors from accessing the SPE's assets in case of bankruptcy. In practice, SPEs were used by banks as a device for circumventing capital adequacy regulations and so reducing the banks' tax bills;

as demonstrated by the Financial Crisis of 2008–2009, the resultant allocation of property rights was not efficient. SPEs and the boundaries that enclose them therefore receive a failing grade when evaluated from an object-based perspective.

Object-based approaches can be understood in moral terms, but they do not lend themselves to a moral analysis of organizational behavior. Because object-based approaches are typically underpinned by efficiency-oriented maxims, they at least implicitly endorse utilitarian views of organizations and of the goals they should pursue; such reductionist approaches can shed light on important questions, but they tell only part of the story. Our motivating example of public attitudes towards Royal Dutch Shell was selected because it clearly demonstrates that, for many people, it feels natural to evaluate an organization as a subject in itself, quite independently of its legal boundary, and not as an object defined by that boundary and used as a tool by human actors. In order to evaluate an organization as a subject, it is necessary to adopt a particular conception of the organization as an entity that can be held morally responsible for its actions. From this perspective, the organization's boundaries separate actions for which it can be held morally responsible from those for which it cannot.

The Organization as a Subject

We have shown how the existing literature on organizational boundaries starts from an understanding of the organization as an object that people use as a tool to achieve a particular goal, decides where the organization stops and the rest of the social world starts in light of that goal, and then uses the resulting insights either to explain observed facts about organizational boundaries, or to suggest ways in which they might usefully be redrawn. We have argued that moral judgments of organizational behavior require a different conception of the organization as a subject; in turn, this requires us to accept that organizations can possess moral agency. In this section, we ask what it means for organizations to be moral agents, and we outline the characterization of moral agency with which we will work.

An organization is a moral agent if it is morally responsible for its actions. Whether or not it is reasonable to view organizational actors as moral agents is contentious. It is clear that an organization cannot possess all of the features of a fully fledged human moral agent; arguments that organizational actors can be morally responsible therefore rest upon the claim that organizations have some important characteristics that are sufficient for an attribution of moral agency. Broadly speaking, analyses of corporate moral agency (i.e., group moral agency) have one of two points of departure. The *individualistic* approach starts by asking what internal characteristics an organization must possess in

order to exhibit metaphysical freedom; if the relevant characteristics are present, then an organization that combines them can justifiably be held morally responsible by external agents. The *relational* approach starts from the observation that people experience emotional responses to corporate actions and then asks if those emotional responses persist when they are scrutinized in light of the organization's internal characteristics.

Individualistic Accounts of Corporate Moral Agency

The individualistic case for moral agency is presented in a broad literature. Contributions to that literature generally identify a moral agent as possessing three characteristics: the capacity to intend states of the world, the ability to act, and autonomy. The first is critical: an agent cannot be held responsible for its actions if it is incapable of forming intentions and, hence, organizations can be corporate moral agents only if they can form intentions that are distinct from their members. A highly influential discussion of this point is due to French (1979), who bases his argument on the Corporate Internal Decision (CID) Structure that organizations use to aggregate information and to arrive at collective decisions. By French's account, a CID structure creates corporate intentions when it synthesizes individual intentions into explicitly organizational goals.

The argument that organizations can form intentions rests upon the formal attribution of intentions that emerge from a CID Structure to the organization within which it sits. This position is sometimes countered by the axiomatic claim that intent has to be the product of independent reasoning. Then, because organizational procedures "cannot create group mental states nor group minds in any literal sense" (Velasquez, 2003, p. 546), and, because intent is claimed to be the product of a single mind, the CID Structure cannot create corporate intentionality. An alternative counterargument focuses on the CID Structure's existence as a system of rules, which cannot have intentions: chess has rules, for example, but only its players have intentions (Keeley, 1981).

The capacity to act is also necessary for an individualistic account of corporate moral agency for, if the organization cannot act, then there is nothing for which to hold it responsible. Of course, all actions in organizational settings are ultimately taken by human actors; individualistic accounts must therefore explain why some actions can nevertheless be attributed to the organization. One such explanation argues that, if an organization creates social conditions that render an action natural and apparently desirable within the organizational context then, because those conditions were created, *someone* will act: the organization thereby causes the action. One achieves an analog in the natural world by bringing a liquid to boiling temperature, at which point it is inevitable that some molecule will have the position and momentum

necessary to crack the surface of the flask that contains the liquid (Pettit, 2007, p. 192). An alternative approach holds that organizations affect the interests and desires of their members; this could happen because of cultural effects (Goodpaster, 1983; Goodpaster & Matthews Jr., 1982), because the shared life experience of members leads them to adopt similar standpoints (Painter-Morland, 2007), or because the employer exerts "affine agency" upon employees who come to view its goals as their own (French, 1994). This type of influence need not bottom out in fully fledged corporate moral agency (Gibson, 2011), but it does provide another reason to believe that organizations can meaningfully be said to exercise agency; some scholars hold that, by virtue of their greater power and resources, corporate agents could even be held to a higher standard than natural persons (Soares, 2003).

Authors who dispute the claim that organizations have the ability to act advance two related arguments. The first maintains that, as legal fictions with no bodily existence, corporations cannot act (Velasquez, 1983). The second holds that individuals are sufficiently autonomous to be unaffected in their deliberations by their immediate social environment: that is, it claims that the (Kantian) freedom of organizational members renders them always entirely responsible for their own actions and, hence, that corporate action is impossible (Kerlin, 1997; Rönnegard, 2015, pp. 40–41).

The final claim of individualistic accounts of corporate moral agency is that, for an organization to be responsible for its actions, it must have chosen the action autonomously. In this context, an agent's action is autonomous if it is justified by beliefs and desires that can be meaningfully attributed to the agent. Following Dennett's (1971) account of intentional systems, it is possible to argue that, for practical purposes, it is sufficient that an attribution of beliefs and desires successfully explains and predicts organizational behavior, regardless of whether those beliefs and desires have a phenomenal reality (Weaver, 1998). However, while some authors claim that organizations cannot have beliefs and desires (Rönnegard & Velasquez, 2017), it is possible to make sense of this type of mental state in an organizational context (Morrison & Mota, 2023). Whether or not they have a phenomenal existence, the beliefs and desires attributed to an organization must be independent of its members if it is to be autonomous. A prominent argument for the possibility of independent organizational beliefs and desires rests on the *discursive dilemma*, which is a mathematical result that considers the aggregation of the beliefs of group members into a logically consistent set of group beliefs (List & Pettit, 2011; Pettit, 2007). Under certain reasonable conditions, every member of the group will disagree with at least some of a logically consistent set of group beliefs: it is in this sense that the group's beliefs and, hence, the resultant actions are autonomous. The organization's formal procedures for

combining individual choices therefore generate autonomous decisions (Pettit, 2007, pp. 180–184).

The discursive dilemma is a mathematical fact and so cannot be contested. But critics of its use have argued that it supports only a very shallow notion of autonomy. Autonomous agents choose their own goals. If they are truly autonomous, then they should be able to understand why they have chosen a goal, and to form higher-order desires about their goals (Frankfurt, 1971). For example, a person who retains racist friendship preferences from childhood should, as a condition of his or her autonomy, be able to desire that he or she shake off those preferences. In contrast, a dog that wants to eat chocolate will never, even after experiencing the adverse effects of eating chocolate, form the higher-order desire that it cease to want to eat chocolate; hence, by this account, a dog is not autonomous.[2] If we accept this understanding of autonomy, then the discursive dilemma is insufficient to establish organizational autonomy, because the dilemma concerns the aggregation of first-order preferences only. Of course, one could argue that the organization's formal deliberation fora allow it to form higher-order desires, but some corporate moral agency skeptics claim that, in fact, this is impossible and, hence, that organizations cannot be autonomous (Rönnegard, 2013; Rönnegard & Velasquez, 2017). A strong statement of this type of position is associated with the perspective that Donaldson (1980) refers to as the "structural restraint view." On this view, the actions of organizations are entirely determined by their structures; this rules out autonomy and, with it, corporate moral agency (Ladd, 1970; Werhane, 1980).

Individualistic accounts of corporate moral agency deploy a procedural understanding of organizations that does not account for important aspects of the way that they are experienced in social interactions. Individualistic accounts start by identifying features of an organization that justify the attribution of moral agency and, on the basis of those features, conclude that it is appropriate to assign blame or praise. But, as a practical matter, this is not how the social processes that underpin the assignment of moral responsibility to organizations work. In practice, when they witness organizational actions, people first experience affective responses such as praise, blame, and outrage, and, as a result of those responses, they then assign moral responsibility

to the relevant organization. Of course, a person's affective response may be mistaken, but that sort of mistake is solved through appropriate reflection. In order to understand these phenomena, we turn in the next section to an alternative approach to moral agency that relies upon the relational experience that actors have of organizations.

Relational Accounts of Corporate Moral Agency

Relational accounts of corporate moral agency start from the practice of holding responsible and ask whether this practice makes sense (e.g., Björnsson & Hess, 2017; Morrison et al., 2022; Shoemaker, 2019; Silver, 2005; Tollefsen, 2003, 2015); this work typically builds upon Strawson's ([1962] 1974) influential account of *reactive attitudes*. Strawson ([1962] 1974, p. 5) characterizes reactive attitudes as the "non-detached attitudes and reactions of people directly involved in transactions with each other"; those responses are naturally experienced in response to others' goodwill or lack of it: love, anger, indignation, and thankfulness are all reactive attitudes (Strawson [1962] 1974, pp. 6–7). By his own account, Strawson's language is "quite unscientific and imprecise" ([1962] 1974, p. 5). Later work builds on Strawson's insights and attempts to be more precise about the definition of a reactive attitude in order to identify a unifying thread that connects the wide range of emotional responses associated with the concept. Mason (2017, p. 155) distinguishes between an "*aretaic*, *appellative*" and a "*deontic*, *imperative* view of the reactive moral sentiments." A deontic reactive attitude is formed in response to the compliance or violation of normative expectations that correspond to interpersonal moral obligations; it is imperative because the underlying expectations are "*addressed as commands to* or *demands of* their targets" (2017, p. 155). In contrast, the expectations that underpin an aretaic reactive attitude are concerned with character ideals but need not concern interpersonal obligations; they are appellative because they express "*appeals* to comport oneself in the manner befitting the ideals at issue" (2017, p. 156). In this article, we are mostly concerned with deontic, imperative reactive attitudes because they are most closely associated with the practice of holding responsible (see also Darwall, 2006; Wallace, 1994) that grounds the approach to corporate moral agency that we deploy in our analysis.

Strawson's account of reactive attitudes is central to his attempt to reconcile moral responsibility with the truth or falsity of the theory of determinism. According to Strawson, the practice of holding responsible makes sense as a response to goodwill or its absence, regardless of whether or not humans possess the sort of metaphysical freedom denied by determinists. By his account, the reactive attitudes are so natural that it is impossible to imagine social life without them: while it is possible to repress the reactive attitudes, it

---

[2] Of course, one can condition a dog not to eat chocolate, or a child not to hit its friends. But the conditioned agent changes its intent in response to physical stimulus (the dog dislikes being sick; the child does not want to be sent to its room). It does not form a reasoned wish that it should adopt a different goal, and then work to achieve that wish; its actions are therefore altered without reference to its will. Pettit (2017, p. 27) characterizes this type of influence as treating the dog or the child "not as a person in the functional sense of the term, but rather as a slave of stimulus and effect."

is unnatural to do so, because it involves treating people as objects, rather than as social interlocutors. Strawson argues that it is precisely because the reactive attitudes are partly constitutive of human nature that it is appropriate to hold other people responsible for their actions. He claims that any attempt to understand the deeper metaphysical reason that we experience a reactive attitude is to "over-intellectualize the facts," which, by Strawson's account, are not subject to external rational justification; instead, "questions of justification are internal to the structure" of reactive attitudes (Strawson [1962] 1974, p. 25).

Although negative reactive attitudes like resentment are a natural human response to the absence of goodwill, there are certain circumstances that can modify or inhibit those attitudes (Strawson, [1962] 1974, pp. 7–10; Wallace, 1994, pp. 118–127; Watson, 1987, pp. 259–261). Strawson divides these circumstances in two groups. The first group comprises situations in which the target of the reactive attitude is fully capable of being held responsible but can appeal to an *excuse* or a *justification* of his or her conduct. Excuses include ignorance, accidents, or the interference of external forces like coercion; justifications correspond to reasons why the agent's conduct might be morally permissible or obligatory. The second group of considerations that can modify or inhibit reactive attitudes concerns an agent's temporary or permanent lack of capacity for moral responsibility, as is the case, for example, with severe cognitive impairment. Such a lack of capacity constitutes an *exemption*; if that exemption is permanent, then the agent cannot take part in moral relationships. Excuses, justifications, and exemptions are evidence that a reactive attitude's internal criteria are not satisfied and, hence, such attitude does not withstand scrutiny.

Relational accounts of corporate moral agency start from the claim that organizations are the sorts of things towards which people form reactive attitudes, and then ask whether those attitudes can survive scrutiny; more specifically, these accounts ask whether there are exempting conditions that would exclude organizations from moral relationships.[3] Hence, relational accounts of corporate moral agency argue that organizations possess the type of normative competence that allows them to be the target of a sustained reactive attitude.

We have now discussed individualistic and relational accounts of corporate moral agency. Both accounts are consistent with the process model that we present in the next section; however, the choice of which account to deploy

matters, because one account may lead to the conclusion that an action sits within an organization's boundary, while the other may lead to the opposite conclusion. In the next subsection, we present the relational account that we use to analyze examples in this article, and we explain why we believe that it may be the more useful approach for the analysis of organizational boundaries.

The Corporate Second-Personal Account of Corporate Moral Agency

When we need to choose an account of corporate moral agency in this article, we use a specific relational account: the *corporate second-personal*, or CSP, account presented by Morrison, Mota, and Wilhelm Jr (2022). In this subsection, we briefly present the CSP account, and we then explain why we believe that it is better suited to practical applications of our theory.

The CSP account builds on Stephen Darwall's (2006) work on the second-person standpoint to conceptualize the competence that a corporate moral agent must exhibit. Darwall (2006) studies the way that agents hold one another to account, and the moral authority that entitles them to do so. He argues that a person who experiences a reactive attitude is entitled to demand a moral account-giving from the responsible actor (see also Wallace, 1994, p. 6). Darwall argues that this type of account-giving appropriately occurs in a dialogue between two people, each of whom recognizes both that the other has moral authority and also that this authority generates reasons to act. When people recognize these facts, they adopt the *second-person standpoint*. Hence, Darwall argues that the right way to demand a moral account is to address one's interlocutor *second personally*. A moral agent is an agent that is *second-person competent*: that is, one that is capable of receiving, deliberating over, and responding appropriately to second-personal addresses. A moral agent must therefore be capable of treating an interlocutor as entitled to respect. In other words, a moral agent must be capable of being motivated by respect in its response to a second-personal address, rather than by some other, possibly instrumental, expedient.

In the relational tradition, the CSP account holds that, for an organization to be a moral agent, it must be possible for people to form reactive attitudes towards the organization. In addition, the CSP account requires those reactive attitudes to survive logical scrutiny so that the organization is not deemed exempted from moral responsibility. That type of exemption occurs when the target of a reactive attitude lacks the "normative competence" to engage in a moral dialogue (Doris, 2002, p. 136; Shoemaker, 2007; Stern, 1974, pp. 77–79; Watson, 1987, pp. 267–268). According to the CSP account, an organization's normative competence derives from the organizational equivalent of the second-personal

---

[3] In this paper, we restrict our analysis of relational accounts of corporate moral agency to the Strawsonian tradition. Other authors have proposed relational approaches to moral responsibility in organizations that build upon different philosophical foundations. For example, Faldetta (2018) discusses theories of responsibility that are grounded on Levinasian ethics.

competence studied by Darwall. That equivalence requires that it be possible to address the organization, and also that it be capable of responding in a way that acknowledges and respects its interlocutor's moral authority. These requirements are captured in three formal *corporate second-personal conditions* (CSP conditions) that an organization must satisfy in order to be a moral agent:

C-1. *Attitude Condition*: If an organization can be held morally responsible, then it must be possible for other agents to form reactive attitudes toward it.
C-2. *Address Condition*: If an organization can be held morally responsible, then it must be possible to trigger a dialogue by addressing reactive attitudes to it.
C-3. *Authority Condition*: If an organization can be held morally responsible, then it must be able to acknowledge the moral authority of other agents. That is,
(1) it must be able to recognize others as valid sources of moral reasons;
(2) it must be able to perform second-personal deliberation;
(3) it must be able to respond appropriately to second-personal addresses.

The CSP account is meta-ethical: it holds that the CSP conditions are necessary and sufficient for corporate moral agency, but it does not claim that any organization actually satisfies them. Some authors argue that at least the Attitude Condition cannot be satisfied, because, by their account, it is impossible to sustain a reactive attitude towards an entity that is incapable of experiencing affect (Sepinwall, 2017; Shoemaker, 2019). In response, the CSP account holds that it is at least possible that an organization could respond adequately to a reactive attitude without experiencing affect (Morrison et al., 2022, pp. 12–14). In contrast to the original statement of the CSP account, which remains agnostic as to the possibility of corporate moral agency in practice (Morrison et al., 2022, p. 323), we assume henceforth that organizations can satisfy the CSP conditions.

For three reasons, we claim that the CSP account is better suited to an analysis of organizational moral boundaries than other accounts of corporate moral agency. First, the CSP account is relational and, as we note above, relational accounts are more naturally aligned with the social processes by which moral agents are held responsible. Second, the CSP account yields a straightforward list of criteria for establishing whether or not an organization is a moral agent and therefore renders the process easier to apply.

The third reason is that, if moral responsibility is assigned under the CSP account, then it is always possible to engage in moral dialogue with the responsible organization. That sort of dialogue is important because it allows for the giving and receiving of reasons that underpins moral learning. Under the CSP account, then, organizations can learn from their moral errors; we view this as an important criterion for membership of the moral community. The CSP account is more demanding than individualistic accounts (Morrison et al., 2022, pp. 341–342): in contrast to this account, an organization that is incapable of moral dialogue could be deemed a moral agent, and so be considered an appropriate target for moral blame, under an individualistic account. The only response to moral blame if the organization cannot engage in moral dialogue is to treat it as an object and to manipulate it accordingly. We view this conclusion as being more aligned with the assignment of legal responsibility, which is concerned with punishment and compensation rather than with the maintenance of appropriate relationships between members of the moral community.

In this section, we have reviewed the literature that studies the organization as a subject and, hence, as a potential moral agent. Broadly speaking, that literature deploys two modes of analysis: an individualistic approach that explains moral agency in terms of the organization's internal characteristics; and a relational approach that focuses on the way that people experience the organization as a social actor. In this article, we adopt the CSP account of corporate moral agency (Morrison et al., 2022), which argues that three corporate second-personal conditions are necessary and sufficient for an organization to possess moral agency. The next section defines the moral boundary of the organization, presents a process for assigning moral responsibility, and uses the CSP account to illustrate that process.

The Moral Boundaries of Organizations

Our analysis of the moral boundaries of organizations asks, first, what sorts of entities have a moral boundary and, second, precisely what that boundary delineates. We have already addressed the first question: an organization can be said to have a moral boundary only if it is a moral agent. The organization could be a legal corporation, a partnership, a community group, or any other type of organized collective. Because we are deliberately agnostic about the structure of the organization, we cannot define its moral boundary in terms of contingent features like its assets or the ways in which its members are organized. What every moral agent has in common is that one can meaningfully hold it morally responsible for its actions, and actions are therefore the relevant unit of analysis in this context.

We answer the second of our questions by defining the moral boundary of a corporate moral agent to encompass the actions that can meaningfully be attributed to that agent. In order to operationalize this definition, we distinguish between *performed* and *performable* actions. Performed actions are easy to identify because they actually happen and, hence, can be observed: for example, an organizational

Fig. 1 Process for the moral attribution of a performed or performable organizational action, OA
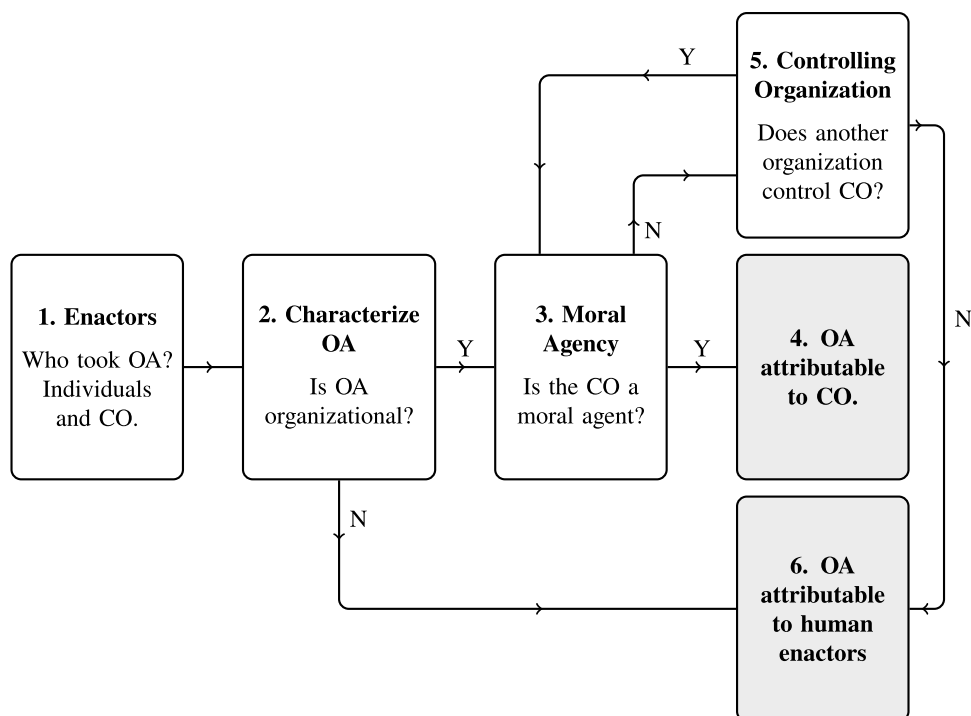
decision to hire someone is documented and promulgated to relevant parties and so can be directly perceived. A performable action is one that could in principle be performed. In order to identify a potential action as performable, one has to know that the organization is capable of performing it. In turn, that requires one to understand what states of affairs the organization might face in the future, the configurations of assets and personnel that would be feasible for the organization in each state of affairs, and what actions the organization could perform with each of those configurations. The imaginative work required to adumbrate every performable action is impossibly high and the moral boundary of the organization therefore cannot be perceived in its entirety; we can only observe the parts of that boundary that emerge when we analyze specific organizational actions, whether performed or performable.

The remainder of this section is organized as follows. We first present a process model for determining whether or not a given organizational action (OA), either performed or performable, lies within an organization's moral boundary. This determination is the first step in the assignment of corporate moral responsibility: performed actions that lie within the moral boundary are actions for which the organization is morally responsible, and moral responsibility is only possible with regard to such actions. In case the organization is not a moral agent, moral responsibility for OA must either be assigned upwards to a controlling corporate moral agent, if it exists, or downwards to specific natural persons. We close the section with a brief discussion of situations in which several corporate moral agents are jointly responsible; those situations include cases in which one corporate moral agent is dependent for its existence upon another.

Unveiling the Moral Boundaries of Organizations

In order to develop our conceptualization of the moral boundaries of organizations, we now present a process model.[4] The model allows us to unveil the moral boundary of an organization by explaining the steps required for the moral attribution of a performed or performable organizational action, OA, to the appropriate agent. The process is illustrated in Fig. 1. We discuss each of the steps in turn.

1. *Enactors*. The first step in the process is to ask who performed the action; that is, who the *enactors* are. This requires the identification of a Candidate Organization (CO) and, possibly, of one or more natural persons who are its members. The CO can be identified according to formal or informal criteria: for example, through legal attribution of OA to the group or through social recognition of OA as attributable to an organization with no legal incorporation. It is possible that more than one CO is identified: one could be the result of legal attribution, and another could be the group agent that is the target of outsiders' reactive attitudes triggered by OA. The identification of more than one CO is unproblematic: first, as will become apparent in the next

---

[4] Our process model is therefore intended to be a part of our theorizing, and not simply to summarize our prior reasoning or the literature that precedes this work.

steps, starting the process with a focal CO does not exclude the consideration of another in later steps; and second, as in the cases of joint corporate moral responsibility discussed below, it may prove necessary to go through the process separately for multiple organizations.

The human enactors may be identified in abstract or specific terms: for example, they can be identified only by role as (say) a company accountant, an HR director, a work team on the factory floor, or a sales department; or they can be identified by name as specific individuals.

2. *Characterize OA*. Step 2 in the process characterizes OA either as being meaningfully organizational, or not. In general, OA is organizational if its human enactors were authorized, instructed, or expected to take OA on behalf of the organization as a result of formal or informal governance mechanisms (Heath & Norman, 2004; Lawson, 2015; Martins, 2018; Morck, 2008; Morrison & Mota, 2023; Westphal & Zajac, 2013). The characterization of OA as organizational, like the assessment of moral agency in step 3, may require specific empirical investigation. If OA can plausibly be characterized as organizational, then the process moves to step 3. If it cannot (for example, because it is clear that organizational members acted against the CO's norms and culture), then moral responsibility for OA is *assigned downwards* to the individual level, as discussed in step 6 below.

3. *Moral Agency*. In order that the CO can have moral boundaries to encompass OA, CO must be a moral agent. Step 3 checks whether this is the case. As discussed in the previous section, our analysis will rely upon a relational theory of corporate moral agency, but the process model could accommodate any perspective. If the CO is a moral agent, then the process concludes at stage 4. If it is not, then the CO has no moral boundaries; in this case, the process moves to step 5.

4. *OA Attributable to CO*. If the CO is a moral agent, then OA lies within its moral boundary.

5. *Controlling Organization*. If CO is not a moral agent, then it is possible that OA lies within the moral boundary of another organization with a particular relationship to the CO: that is, the moral boundary of a controlling organization may enclose the actions performed by the members of the CO. Controlling organizations stand in legal or de facto relationships of ownership and control to the CO. If a controlling organization exists, then it becomes the new CO and the process returns to step 3. In case the controlling organization is a moral agent, OA lies within its moral boundary, and we say moral responsibility for OA is *assigned upwards*; we present an example below. If there is no controlling organization, then the locus of moral responsibility for OA resides at the individual level, as discussed in the next step.

6. *OA Attributable to Human Enactors*. If no organization can be held morally responsible for OA, then OA's human enactors are individually or jointly responsible; as in step 2,

we say that moral responsibility for OA is *assigned downwards*. This step requires the human enactors to be characterized in specific terms; if that type of characterization did not occur in step 1, then additional work is required in this step. Precisely how individual responsibility is apportioned when there is more than one enactor is outside the scope of this article.

To illustrate how this process could be applied in practice, we consider a financial services firm, "Iron Bank," that wishes to offer complex long-maturity currency derivative trades to institutional investors.[5] Iron Bank's clients will refuse to deal derivative contracts with any firm that has a very high level of debt. Iron Bank could reduce its indebtedness, but doing so would have very costly tax implications. So, instead, Iron Bank sets up Syrio Forex, a separately capitalized company, to deal with its institutional clients. Syrio is a Special Purpose Entity, or SPE: legally, it is distinct from Iron Bank and, hence, is not dragged down in the event of Iron Bank's bankruptcy. Syrio is nevertheless entirely owned by Iron Bank, which staffs it, sets its strategy, and controls its operations.

Consider a derivatives transaction T that is formalized in a legal agreement between Syrio and an institutional investor. T is a performed organizational action and, hence, the process of Fig. 1 can be used to establish to which agent T should be morally attributed. Syrio is a Candidate Organization (CO), and the human enactors, characterized in abstract terms, are Syrio's authorized signing officers (step 1). Because the signing officers are authorized to perform T on behalf of Syrio and T is covered by an earlier master agreement between Syrio and the institutional investor, T can meaningfully be understood as organizational (step 2).

T is legally attributable to Syrio, but moral attribution is not possible because Syrio is not a moral agent (step 3). In order to demonstrate this, we rely upon the CSP account of corporate moral agency and check whether Syrio satisfies the CSP conditions. The Attitude Condition is violated because Syrio's clients view themselves as dealing with Iron Bank, even though Iron Bank is not a formal party to the transaction; in fact, Syrio is virtually invisible and, hence, is not the target of reactive attitudes. Syrio cannot satisfy the Address Condition because any messages addressed to Syrio are considered, and responded to, by Iron Bank, so that it is impossible to trigger a dialogue with Syrio. Finally, Syrio is incapable of recognizing the moral authority of other

---

[5] Here, and elsewhere in the article, we follow the philosophical tradition of motivating formal arguments with fictional vignettes, or "thought experiments" (see, e.g., Ichikawa and Jarvis 2009; Williamson 2005). This approach allows us to isolate the key features that generate moral intuitions in a given situation while avoiding the conflating effects and cultural baggage associated with real-world examples with which readers may be familiar.

agents because it does not have standalone governance systems and so is incapable of second-personal deliberation; consequently, Syrio cannot satisfy the three requirements of the Authority Condition. Although Syrio is not a moral agent, Iron Bank, which owns and controls it, could be; Iron Bank therefore becomes the new CO (step 5). If Iron Bank satisfies the CSP Conditions and therefore is a moral agent (step 3), then its moral boundary encompasses the actions that are legally attributable to Syrio. In this case, moral responsibility is *assigned upwards* to Iron Bank (step 4). This is because, first, any reactive attitudes that are triggered by actions that are legally attributed to Syrio are directed towards Iron Bank; and, second, Iron Bank is the entity that receives, processes, and responds to second-personal addresses of those reactive attitudes. This example makes it clear that a firm's moral boundary need not coincide with its legal boundary.[6]

If Iron Bank does not satisfy the CSP Conditions (step 3), and is not controlled by any other group agent (step 5), then responsibility must be *assigned downwards* to individual enactors (step 6). Moral responsibility for T may not be restricted to the authorized signing officers; other members of both Syrio and Iron Bank may be jointly morally responsible for the transaction.

To illustrate an alternative path through the process, consider a situation in which Thyco, a derivatives trader working for Syrio (step 1), punches a trader from Braavobank (Iron Bank's biggest competitor) during an industry drinks event. Thyco is formally employed by Syrio, which authorized his presence at the industry event; however, it is very unlikely that such authorization extended to physical aggression. It follows that the punch cannot be characterized as organizational (step 2), and moral responsibility for the aggression must therefore be assigned downwards to Tycho (step 6). Note that downwards assignment does not exclude the possibility that Syrio or Iron Bank bears moral responsibility for authorizing Tycho to attend the event if the aggression was foreseeable: although it is related, the authorization itself is a separate action that requires separate assessment.

Joint Responsibility and Nested Moral Agency

When a corporate moral agent is held morally responsible for an action, it may share that responsibility with other moral agents, either human or organizational. As mentioned above, that humans and corporate moral agents can be jointly responsible is fairly uncontroversial (see, e.g., Constantinescu & Kaptein, 2021; Hess, 2017; Mathiesen, 2006). Similarly, if two corporate moral agents contributed to an action that falls within the moral boundaries of both, then they can be held jointly responsible for it. In general, joint responsibility in the organizational context raises questions regarding the apportionment of responsibility precisely as it does when responsibility is shared by individual agents; those questions are outside the scope of this article.

We now turn to a discussion of *nested moral agency*, which has no analogue for human actors, and which sheds new light on questions of joint moral responsibility. Nested moral agency arises when some of the members of a corporate moral agent (a *host*) form a separate but dependent group that is also a corporate moral agent (a *nested moral agent*).

In order to show how a nested moral agent can emerge, consider a variant of our Iron Bank example in which Iron Bank satisfies the CSP conditions, and, hence, is a moral agent. Suppose that a number of Iron Bank's members combine to form an Employee Resource Group (ERG) in order to create a peer-support network for LGBTQ+ staff. Over time, the ERG may begin actively to lobby for more inclusive organizational policies, to engage with local non-governmental organizations, and to collaborate with civil society organizations to foster diversity and inclusion in the local community. In performing these activities, the ERG becomes visible both to Iron Bank and to outsiders. In order to operate effectively, the ERG may also develop governance procedures that facilitate dialogue with outsiders; if, in addition, those governance procedures enable the ERG to recognize the moral authority of its interlocutors, then the CSP conditions are satisfied. When this happens, the ERG becomes a nested moral agent whose host is Iron Bank.

An action taken by an individual member of a nested moral agent could be attributable to that individual, the host, the nested moral agent, or any combination of the three. The process outlined above should therefore be followed separately for the host and the nested agent. Because both satisfy the CSP conditions (step 3), the question of whether the action can be understood as organizational (step 2) is critical.

Suppose that the ERG has an existing commitment to a partnership with a local civil society organization, Safe Space, and that the ERG informs Safe Space of its intention to withdraw from the partnership and, hence, to stop financing the associated projects. We present three scenarios in which the withdrawal is organizational for, and so morally attributed to, the ERG, its host Iron Bank, and both. First, imagine that the ERG's withdrawal decision was taken as a result of its own reporting systems and governance

---

[6] As discussed in Footnote 5, we use the fictional example of Iron Bank in order to abstract away from complicating factors that may obscure the development of moral intuition about real-world cases. But such real-world cases abound: firms commonly use Special Purpose Entities to separate complex trades from the rest of their business. See, for example, Gorton and Souleles (2007) for a discussion of SPEs and financial institutions and Sainati et al. (2020) for a discussion of SPEs in infrastructure megaprojects.

structures, without the involvement of Iron Bank. In this case, the withdrawal is morally attributable to the ERG. Second, imagine that Iron Bank ordered the ERG to end the partnership; in this case, the withdrawal is morally attributable to Iron Bank. Third, suppose that the decision to end the partnership was the result of an agreement between Iron Bank and the ERG that Safe Space was an inappropriate partner because it did not further Iron Bank's commercial ends. In that case both Iron Bank and the ERG contributed to the decision, and it is therefore attributable to both.

Creating, Destroying, and Changing the Moral Boundaries of Organizations

In this section, we show how an organization could create, destroy, or change its moral boundary. An organization can possess a moral boundary if and only if it is a moral agent; the acquisition and loss of a moral boundary can be the result of deliberate actions either inside or outside the organization. The notion of a performable action is central to our analysis of changes to the moral boundary of a moral agent: the boundary expands if the agent's range of performable actions broadens, and it contracts if that range narrows. We will use the CSP account of corporate moral agency throughout our discussion of the creation, destruction, and modification of moral boundaries but, as in the previous section, the analysis could be founded upon a different account of corporate moral agency.

Creating Moral Boundaries

An organization that is not a moral agent can decide to change itself so as to become one. Under the CSP account, this is only possible in case the organization already possesses some ability to deliberate and to establish its own goals and procedures. To illustrate the acquisition of moral agency, consider a secret society. The society does not satisfy the CSP conditions, because it is completely invisible, and therefore cannot be the target of reactive attitudes; equally, the society cannot be addressed, and cannot speak to outsiders (Morrison et al., 2022, p. 341). However, if the society's formal and informal procedures allow it to redefine its presentation to the outside world, its actions, and its motivations, then it can change itself so as to satisfy the CSP conditions. The secret society can choose to make itself visible (for example, by creating and developing a brand), to create or facilitate fora to enable dialogue triggered by reactive attitudes, and to ensure that it deliberates second personally. In doing so, it becomes a moral agent, and it does so by choice. Equally, a firm that satisfies the Address and Attitude Conditions, but whose culture prevents it from satisfying the Authority Condition, can attempt to change its culture so as to become a moral agent; although it is difficult to change organizational culture, it is not impossible to do so (Payne et al., 2022; Wines & Hamilton, 2009). Regardless of the path an organization takes to moral agency, getting there implies the creation of a moral boundary.

Destroying Moral Boundaries

We now discuss situations in which moral boundaries are destroyed as a result of the loss of moral agency. Consider an organization that satisfies the CSP conditions and, hence, possesses moral agency. Suppose that, over time, the chief executive officer (CEO) becomes entrenched. By this we mean that it becomes impossible to dislodge the CEO, that he or she controls the appointment of board members, and that any decisions taken by the board are really dictated by the CEO. When this happens, it is possible that outsiders begin to form reactive attitudes towards the CEO, rather than the organization; if that happens, the organization no longer satisfies the Attitude Condition. Equally, if the organization is entirely controlled by the CEO, then any fora for dialogue with outsiders are unlikely to be meaningful, and the organization loses the ability to respond appropriately to second-personal addresses: any response would be the CEO's, and not the organization's. In short, the entrenched CEO undermines the CSP conditions to such an extent that the organization can no longer be viewed as a moral agent. Of course, the precise time at which this happens may be hard for outside observers to identify: the transition between the morally responsible organization and the organization whose actions can only be morally attributable to the CEO may not be observable. But the destruction of the moral boundary nevertheless occurs, and it entails downwards assignment of responsibility.

An organization can also cease to be a moral agent as a result of mergers or acquisitions. To fix language, suppose that organization A acquires organization B, and that A and B are moral agents. After the acquisition, A must decide whether A and B should have the same brand, and whether or not to merge the governance systems of A and B. These integration decisions may affect B's moral agency. Suppose first that A elects to be an arm's-length investor. This happens, for example, if A is a private equity firm that adds B to its portfolio. In that case, B retains its own brand, its own board of directors, and its own communications function. As a result, outsiders continue to form reactive attitudes towards B, and the systems that allow B to receive and to respond to second-personal addresses continue to operate: consequently, B retains its moral agency after the acquisition. If, on the other hand, A is an organization that opts to merge its brand with B and to assume direct control over B's assets and operating decisions, then B's moral boundary is destroyed and its moral agency dissolves into A's. This

happens because it is no longer possible to form reactive attitudes specifically towards B; because it is only possible to address reactive attitudes to the merged organization; and because all deliberation over those addresses is performed collectively by the merged organization.

Changing Moral Boundaries

A corporate moral agent alters its moral boundary when it alters its range of performable actions. It is easier for a corporate moral agent to accomplish this than it would be for a human person, because an organization can very rapidly change its personnel or its structure in a way that increases or reduces the actions that it is able to perform. In contrast, a human person typically has to invest significant time and effort to acquire the new skills it would need to learn a new action, and humans un-learn an ability to perform a given action only slowly, and sometimes with great difficulty. In order to illustrate how a corporate moral agent might change its moral boundaries, we return to our Iron Bank example. In this subsection, we will assume that Iron Bank is a moral agent.

Suppose that Iron Bank realizes that it has an obligation to conduct human rights due diligence, and that this is not one of its performable actions. In order that Iron Bank acquires the capacity to discharge this obligation, it must change its structure and possibly its personnel. The structural change would ensure that relevant actions were induced, authorized, or expected by Iron Bank whenever human rights were at stake; such actions would then be meaningfully organizational. The structural change might involve setting up a new department, creating new management positions, or changing the official responsibilities of existing departments. Structural change alone is unlikely to suffice, because human rights due diligence is a complicated and specialized activity that Iron Bank's existing employees may not be able to perform. Hence, Iron Bank may have to train existing staff or to hire new ones; it could also acquire another bank with the structure and capacity to perform human rights due diligence. However Iron Bank gets there, when it acquires the ability to perform the associated actions, it expands its moral boundary.

Suppose now that Iron Bank decides on moral grounds that it no longer wishes to market its derivative products to retail clients and, hence, that it wishes to remove that sort of transaction from its list of performable actions. In order to do so, it would change its structure so as to stop supporting retail marketing by removing authorizations and excluding relevant positions from its organizational chart. It could fire staff involved in retail marketing, or it could redeploy them; it could also attempt to spin off its retail activities as a separate standalone business. Any of these organizational changes would contract Iron Bank's moral boundary.

In some circumstances, an organization might change in a way that renders some actions unperformable and simultaneously introduces new performable actions. That sort of change alters the organization's moral boundary, but the change need neither expand nor contract that boundary.

Conclusion

Any discussion of corporate responsibility necessarily involves an understanding of who the responsible agent is, and what actions can be attributed to it. This article presents a theory of the moral boundaries of organizations that provides a path to that understanding. In order to construct our argument, we introduce a distinction between the organization as an object and the organization as a subject: thinking of the organization as a subject requires us to focus on moral agency and also releases us from legal and economic understandings of organizations.

Prior discussions of where organizations end and the rest of the social world begins are dominated by the vast literature that was seeded by Coase's seminal (1937) analysis of organizational boundaries. Because that literature is concerned with organizations as tools that get things done—that is, with organizations as objects—it is ill-suited for our purposes: if we attempt to think about moral responsibility in Coasean terms, then we risk assigning that responsibility incorrectly and unfairly. Our first contribution is to explain why this is the case, and clearly to distinguish between the legal and moral boundaries of organizations.

Precisely because, in our telling, organizations are subjects, their boundaries are expressed in terms of the actions that they can perform, rather than their legal or institutional features. Our second contribution is to explain what this means, and to offer a process model that allows us to determine whether a given action lies within an organization's moral boundary.

The organizations that we encounter in day-to-day life are frequently very complex: for example, many multinational enterprises have multiple subsidiaries that, in turn, own and control other organizations. Consequently, when we attempt to assign moral responsibility for a given action, we are often faced with a network of interconnected organizations. Our third contribution is to present an analytic framework that makes these situations more tractable by explaining when moral responsibility should be assigned to the various elements of such a network. Our framework clarifies situations in which one organization shares responsibility with another or, as is more frequently discussed, with humans; it also accounts for cases in which, notwithstanding the organizational apparatus surrounding an action, responsibility devolves solely upon one or more human agents.

Our analysis of joint moral responsibility allows us to understand more clearly how, unlike humans, corporate moral agents can be nested. As far as we are aware, ours is the first article to analyze this type of corporate moral existence; doing so is the fourth contribution of this article. Realizing that one corporate moral agent can host another reveals more complex and nuanced forms of moral responsibility: a person who lacks the concept of nested moral agency may incorrectly and unfairly attribute the actions of a nested agent to its host.

In addition to extending the academic literature on corporate moral agency and corporate moral responsibility, our work provides insights of relevance to practitioners and policy-makers. Because our framework allows for clear attribution of moral responsibility and highlights the ways in which moral boundaries can be affected by organizational choices, it has important implications for corporate governance: specifically, our work may inform the design of reporting lines, the management of organizational culture, and the ways that organizations communicate with external stakeholders. Our work could also inform decisions about corporate structure, such as the decision to expand via branches or subsidiaries. Finally, our theorization contributes to complex policy debates about the allocation of legal, as opposed to moral, responsibility among human and corporate actors. Even if one believes that legal and moral responsibility need not (or even should not) fully coincide, it is important that legal systems develop in a way that is at least consistent with the proper attribution of moral responsibility to organizations and their members; new insights on moral responsibility can inform legal developments in this area.

Our analysis reveals several questions that merit further investigation. First, our work sheds light upon a class of moral obligations that has received minimal prior attention: namely, obligations that attach to the choices that a corporate moral agent makes about its structure. For example, suppose that a parent organization with moral agency sets up an SPE in order to avoid legal responsibility for a performable action. Then, by our account, the SPE's actions fall within the moral boundary of the parent organization. But the choice to create an SPE was itself an action that might generate new moral obligations for the parent related to the use that it makes of its new corporate structure, including situations in which the parent might try to blame the SPE in order to escape public accountability.

Our work also suggests a second, empirical, set of questions about how organizational choices that affect the moral boundary are made and how they are experienced by organizational stakeholders. For example, future work could investigate the extent to which organizational decision makers consciously consider the moral implications of creating a nested moral agent, a subsidiary that has moral agency, or a bankruptcy-remote entity that has no moral agency.

A third set of questions relates to the phenomenon of nested moral agency that we identify in our analysis. Future work could study the processes by which this type of agency develops in practice; which moral boundary people perceive as relevant when dealing with a nested moral agent; how people direct reactive attitudes (say, of praise or blame) towards nested agents; and how those reactive attitudes develop as people come better to understand those agents' structures.

Declarations

References

Araujo, L., Dubois, A., & Gadde, L.-E. (2003). The multiple boundaries of the firm. *Journal of Management Studies, 40*(5), 1255–1277.

Armour, J., Hansmann, H., & Kraakman, R. (2017). Agency problems and legal strategies. In *The anatomy of corporate law: A comparative and functional approach* (3rd ed., pp. 29–47). Oxford, UK: Oxford University Press.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Barney, J., Wright, M., & Ketchen, D. J., Jr. (2001). The resource-based view of the firm: 10 years after 1991. *Journal of Management, 27*(6), 625–641.

Björnsson, G., & Hess, K. (2017). Corporate crocodile tears? On the reactive attitudes of corporate agents. *Philosophy and Phenomenological Research, 94*(2), 273–298.

Coase, R. H. (1937). The nature of the firm. *Economica, 4*(16), 386–405.

Constantinescu, M., & Kaptein, M. (2021). Virtue and virtuousness in organizations: Guidelines for ascribing individual and organizational moral responsibility. *Business Ethics, the Environment & Responsibility, 30*(4), 801–817. https://doi.org/10.1111/beer.12373

Darwall, S. L. (2006). *The second-person standpoint: Morality, respect, and accountability*. Cambridge, MA: Harvard University Press.

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy, 68*(4), 87–106.

Donaldson, T. (1980). Moral agency and corporations. *Philosophy in Context, 10*, 54–70.

Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge: Cambridge University Press.

Faldetta, G. (2018). A relational approach to responsibility in organizations: The logic of gift and Levinasian ethics for a 'corporeal' responsibility. *Culture and Organization, 24*(3), 196–220.

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy, 68*(1), 5–20.

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

French, P. A. (1994). *Corporate ethics*. New York, NY: Wadsworth.

Gibson, K. (2011). Toward an intermediate position on corporate moral personhood. *Journal of Business Ethics, 101*, 71–81.

Goodpaster, K. E. (1983). The concept of corporate responsibility. *Journal of Business Ethics, 2*(1), 1–22.

Goodpaster, K. E., & Matthews, J. B., Jr. (1982). Can a corporation have a conscience? *Harvard Business Review, 60*, 132–141.

Gorton, G. B, & Souleles, N. S. (2007). Special purpose vehicles and securitization. In *The risks of financial institutions* (pp. 549–602). Chicago: University of Chicago Press.

Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy, 94*(4), 691–719.

Hansmann, H., & Kraakman, R. (2000). The essential role of organizational law. *Yale Law Journal, 110*(3), 387–440.

Hansmann, H., Kraakman, R., & Squire, R. (2005). Law and the rise of the firm. *Harvard Law Review, 119*, 1335–1403.

Hart, O. D., & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy, 98*(6), 1119–1158.

Heath, J., & Norman, W. (2004). Stakeholder theory, corporate governance and public management: What can the history of state-run enterprises teach us in the post-Enron era? *Journal of Business Ethics, 53*(3), 247–265.

Hess, K. (2017). The unrecognized consensus about firm moral responsibility. In E. W. Orts & N. Craig Smith (Eds.), *The moral responsibility of firms.* Oxford: Oxford University Press.

Ichikawa, J., & Jarvis, B. (2009). Thought-experiment intuitions and truth in fiction. *Philosophical Studies, 142*(2), 221–246.

Kast, F. E., & Rosenzweig, J. E. (1972). General systems theory: Applications for organization and management. *Academy of Management Journal, 15*(4), 447–465.

Keeley, M. (1981). Organizations as non-persons. *The Journal of Value Inquiry, 15*(2), 149–155.

Kerlin, M. J. (1997). Peter French, corporate ethics and the wizard of oz. *Journal of Business Ethics, 16*, 1431–1438.

Ladd, J. (1970). Morality and the ideal of rationality in formal organizations. *The Monist, 54*(4), 488–516.

Lawson, T. (2015). The nature of the firm and peculiarities of the corporation. *Cambridge Journal of Economics, 39*(1), 1–32.

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

Mahoney, J. T., & Rajendran Pandian, J. (1992). The resource-based view within the conversation of strategic management. *Strategic Management Journal, 13*(5), 363–380.

Martins, N. O. (2018). Justice and the social ontology of the corporation. *Journal of Business Ethics, 153*(1), 17–28.

Mason, M. (2017). Reactive attitudes and second-personal address. In R. Debes & K. R. Stueber (Eds.), *Ethical sentimentalism: New perspectives* (pp. 133–152). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316105672.009

Mathiesen, K. (2006). We're all in this together: Responsibility of collective agents and their members. *Midwest Studies in Philosophy, 30*, 240–255.

Morck, R. (2008). Behavioral finance in corporate governance: Economics and ethics of the devil's advocate. *Journal of Management & Governance, 12*(2), 179–200.

Morrison, A. D., & Mota, R. (2023). A theory of organizational purpose. *Academy of Management Review, 48*(2), 203–219. https://doi.org/10.5465/amr.2019.0307

Morrison, A. D., Mota, R., & Wilhelm Jr, W. J. (2022). Relationships, authority, and reasons: A second-personal account of corporate moral agency. *Business Ethics Quarterly, 32*(2), 322–347.

Painter-Morland, M. (2007). Defining accountability in a network society. *Business Ethics Quarterly, 17*(3), 515–534.

Payne, D., Trumbach, C., & Soharu, R. (2022). The values change management cycle: Ethical change management. *Journal of Business Ethics, 188*, 429–440. https://doi.org/10.1007/s10551-022-05306-8

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

Pettit, P. (2017). The conversable, responsible corporation. In E. W. Orts & N. Craig Smith (Eds.), *The moral responsibility of firms.* Oxford: Oxford University Press.

Reed, S. (2021). Britain's supreme court says shell can be sued over spills. *New York Times*, 13 February 2021, sec. B, 4.

Robertson, D. H. (1923). *The control of industry*. London, UK: Harcourt, Brace.

Rönnegard, D. (2013). How autonomy alone debunks corporate moral agency. *Business & Professional Ethics Journal, 32*(1/2), 77–107.

Rönnegard, D. (2015). *The fallacy of corporate moral agency*. Dordrecht: Springer.

Rönnegard, D., & Velasquez, M. (2017). On (not) attributing moral responsibility to organizations. In E. W. Orts & N. Craig Smith (Eds.), *The moral responsibility of firms.* Oxford: Oxford University Press.

Sabel, C. F. (1993). Studied trust: Building new forms of cooperation in a volatile economy. *Human Relations, 46*(9), 1133–1170.

Sainati, T., Locatelli, G., Smith, N., Brookes, N., & Olver, G. (2020). Types and functions of special purpose vehicles in infrastructure megaprojects. *International Journal of Project Management, 38*(5), 243–255.

Sepinwall, A. J. (2017). Blame, emotion, and the corporation. In E. W. Orts & N. Craig Smith (Eds.), *The moral responsibility of firms* (pp. 143–66). Oxford: Oxford University Press.

Shoemaker, D. (2007). Moral address, moral responsibility, and the boundaries of the moral community. *Ethics, 118*(1), 70–108.

Shoemaker, D. (2019). Blameworthy but unblamable: A paradox of corporate responsibility. *Georgetown Journal of Law and Public Policy, 17*, 897.

Silver, D. (2005). A Strawsonian defense of corporate moral responsibility. *American Philosophical Quarterly, 42*(4), 279–293.

Soares, C. (2003). Corporate versus individual moral responsibility. *Journal of Business Ethics, 46*, 143–150.

Stern, L. (1974). Freedom, blame, and moral community. *The Journal of Philosophy, 71*(3), 72–84.

Strawson, P. F. ([1962] 1974). Freedom and resentment. In *Freedom and resentment and other essays* (pp. 1–28). Oxford, UK: Oxford University Press.

Tollefsen, D. P. (2003). Participant reactive attitudes and collective responsibility. *Philosophical Explorations, 6*(3), 218–234.

Tollefsen, D. P. (2015). *Groups as agents*. New York: Wiley.

Velasquez, M. G. (1983). Why corporations are not morally responsible for anything they do. *Business & Professional Ethics Journal, 2*(3), 1–18.

Velasquez, M. G. (2003). Debunking corporate moral responsibility. *Business Ethics Quarterly, 13*(4), 531–562.

Wallace, R. J. (1994). *Responsibility and the moral sentiments*. Cambridge, MA: Harvard University Press.

Watson, G. (1987). Responsibility and the limits of evil—Variations on a Strawsonian theme. In F. Schoeman (Ed.), *Responsibility, character, and the emotions: New essays in moral psychology*. Cambridge: Cambridge University Press.

Weaver, W. G. (1998). Corporations as intentional systems. *Journal of Business Ethics, 17*, 87–97.

Werhane, P. H. (1980). Formal organizations, economic freedom and moral agency. *Journal of Value Inquiry, 14*, 43–50.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180.

Westphal, J. D., & Zajac, E. J. (2013). A behavioral theory of corporate governance: Explicating the mechanisms of socially situated and socially constituted agency. *Academy of Management Annals, 7*(1), 607–661. https://doi.org/10.5465/19416520.2013.783669

Williamson, O. E. (1975). *Markets and hierarchies*. New York, NY: Free Press.

Williamson, O. E. (1985). *The economic institutions of capitalism*. New York, NY: Free Press.

Williamson, T. (2005). The presidential address: Armchair philosophy, metaphysical modality and counterfactual thinking. *Proceedings of the Aristotelian Society, 105*(1), 1–23. https://doi.org/10.1111/j.0066-7373.2004.00100.x

Wines, W. A., & Hamilton, J. B. (2009). On changing organizational cultures by injecting new ideologies: The power of stories. *Journal of Business Ethics, 89*(3), 433–447. https://doi.org/10.1007/s10551-008-0009-5

Womack, J. P., Jones, D. T., & Roos, D. (2007). *The machine that changed the world: The story of lean production—Toyota's secret weapon in the global car wars that is now revolutionizing world industry*. New York: Simon and Schuster.