

How Should We Interpret Institutional Duty-Claims?

Christoffer S. Lammer-Heindel

Abstract

It is rather natural to suppose that what we mean when we say that an institutional organization has a moral duty is parallel to whatever it is that we mean when we say that an individual has a duty. I challenge this interpretation on the grounds that it assumes that institutional organizations possess those characteristics or abilities requisite for moral agency—an assumption which I argue is highly suspicious. Against such an interpretation, I argue that we have very good reasons to suppose that the term ‘has a duty’ is used equivocally across individual and institutional contexts. In other words, the meaning of an institutional duty-claim is quite different than that of an individual duty-claim, so much so that we ought to recognize that institutional duty-claims are not really duty-claims at all.

Keywords

corporate personhood, moral duties, institutional obligations

1. Introduction

It is rather commonplace to speak about institutional organizations as though they have moral duties and obligations. For example, in the context of an examination of payments made by BP to Gulf Coast residents following the Deep Water Horizon Oil Spill, the United States Congress’s Congressional Research Service explained that in addition to “a legal obligation under the Oil Pollution Act” to issue payments to affected residents, “BP has a moral duty stemming from its responsibility for the spill” (Sherlock, et al., 2010, p. 5). And, writing for the *Financial Times* about the implications of recent changes to the economy, Laurence Fink (2012) has argued, “Companies have a moral responsibility to help both full-time and part-time employees to save enough.”

What, precisely, do we mean when we make such claims? How should we analyze such statements? The natural way to answer these questions is to clarify what we mean when we say that an *individual* has a duty to act in a certain way and to draw a parallel between individual and institutional duty-claims. After all, it is apparent that an *individual* duty-claim such as, “John Doe has a moral duty to compensate the victims of his accident,” and the *institutional* duty-claim, “BP has a moral duty to compensate the victims of its oil spill,” share the same basic form: a subject *S* is said to have a duty to do *a*. Presumably, then, what we mean when we say that an institutional organization has a duty parallels whatever it is that we mean when we say that an individual person has a duty.

This, we might suppose, is a plausible candidate for a “commonsense” interpretation of institutional duty-claims, at least insofar as it appears to be fairly straightforward and unsophisticated. Moreover, it preserves the sense in which the term ‘has a duty’ and its cognates are used univocally, rather than equivocally, across both individual and institutional contexts.

In spite of its initial appeal, I wish to challenge this interpretation by showing that it is, in the end, an interpretation that carries with it some highly sophisticated

theoretical baggage—baggage which I think we ought to leave behind. Against the straightforward interpretation, I will suggest that we have very good reasons to suppose that the term ‘has a duty’ is used equivocally across individual and institutional contexts. More precisely, the meaning of an institutional duty-claim is quite different than that of an individual duty-claim, so much so that we ought to recognize that institutional duty-claims are not really *duty*-claims at all.

2. An Analysis of Individual Duty-Claims

The interpretation of institutional duty-claims which I shall be criticizing supposes that they are strictly parallel to individual duty-claims. Let us, then, briefly examine what we mean when we say of an *individual* that she has a duty.

Etymologically, the term ‘duty’ is associated with the Latin term for debt; thus, what one has a duty to do may be, and often has been, understood to be a function of what one owes to someone, or what is proper given one’s station or position. Duty-claims are thus a kind of normative claim: to say that one has a duty to do some action *a* or pursue some goal *g* is at least to say that one is *justified* in doing *a* or pursuing *g*—i.e., one has a justifying reason to do *a* or pursue *g*.

Of course, when a claim of the form, “*S* has a duty to *a*” is interpreted as expressing the fact that *S* has a justifying reason to *a*, it would not be incoherent or contradictory to go on and admit that *S* also has justifying reasons to abstain from or omit doing *a*. This is owing to the fact that a person can, presumably, have a variety of reasons for acting in a variety of ways, not all of which can in fact be acted upon and some of which may conflict with one another. In his influential book, *The Right and the Good*, W.D. Ross usefully distinguished between what he called “*prima facie duties*” and one’s “*duty sans phrase*.”

I suggest ‘*prima facie duty*’ or ‘*conditional duty*’ as a brief way of referring to the characteristic (quite distinct from that of being a duty proper) which an act has, in virtue of being of a certain kind (e.g. the keeping of a promise), of

being an act which would be a duty proper if it were not at the same time of another kind which is morally significant. Whether an act is a duty proper or actual duty depends on all the morally significant kinds it is an instance of. (2002, pp. 19-20)

Ross goes on to explain that a *prima facie* duty is not itself a duty in the proper sense, “but something related in a special way to duty” (2002, p. 20). The way in which *prima facie* duties are related to duties in the proper sense is that *prima facie* duties are simply reasons for action which figure into deliberation as input, whereas one’s duty proper is whatever is supported by the preponderance of these reasons—i.e., the output of moral deliberation is what we take to be our duty *sans phrase*. Whether a particular *prima facie* duty is also one’s duty proper will depend on whether there are stronger, conflicting reasons to act in some other way; if there are not, then this *prima facie* duty would thereby also be one’s duty proper. Ross is rightly inclined to reserve the term ‘duty’ to denote that which is supported by the preponderance of reason—that course of action which, all things considered, is what one ought to do. Thus, when we say of a person that she has a duty to do something, we ought to take ourselves to mean that she is required or bound to act in that way; more precisely, she is justified in undertaking the action and she lacks sufficient justification for failing to undertake the action. If one is merely justified in doing *a*, but not unjustified in failing to do *a*, then one does not have duty to do *a*. In such a case, *a* would be merely optional, not what is owed to another and not what is *proper* given one’s station.

Before moving on, let me make two additional important points. First, when we say, “*S* has a duty to *a*,” we imply that *S* has (at least in principle) access to some consideration which would justify *S* doing *a*. If a person cannot access the consideration which counts in favor of her acting, quite obviously she does not have a reason, does not have a duty to act, and hence the allegation that she does would be erroneous, or false. To be sure, a third party may have reasons to suppose that *S*’s doing *a* would be justified. For example, it may be good (in the sense of being productive of value) if *S* did *a*. But if *S* does not have epistemic access to the fact that her doing *a* would be justified in some sense or another, she cannot, strictly speaking, be said to possess a justifying reason to *a*. This reveals, at least in part, why we do not think that non-human animals, infants, or individuals with profound mental disabilities have duties or obligations: we are not confident that they have access to reasons which would justify action; they cannot comprehend any reasons they might otherwise be imagined to have for acting in particular ways, or at least we do not take ourselves to be justified in supposing that they do. The point here is that duty-claims are not to be analyzed merely in terms of reasons (justification) available for action, but rather in terms of reasons (justification) which the subject has (or at least epistemically ought to have) access to. This preserves the sense in which when we say of a person that she has a duty to do something or another we mean to suggest that she is in some sense required or bound to act in that way; she is bound by reason—that is, she is bound by the reasons which she in fact has or can reasonably be expected to have.

The second point is related to the first by way of implication. It is simply this: there is a significant difference between the claim, “It would be good if *S* did *a*,” and the claim, “*S* has a duty to *a*.” While the latter is a duty-claim, the former is what I shall call a value-claim: it expresses the notion that a certain state of affairs—perhaps the mere doing of *a* on *S*’s part, or perhaps the consequences of *S* doing *a*—would be valuable in some sense. The latter claim is not a claim about the value of a certain state of affairs. It is, rather, a claim about the person *S*—namely,

what *S* has reasons (justification) to do. Significantly, non-moral agents such as toddlers and non-moral entities such as water heaters and cars can be the subjects of a value-claim, so long as there is some relevant sense in which these beings or entities can “do” things. However, only a rational agent can be the proper subject of a duty-claim.

3. Problematizing Institutional Duty-Claims

It follows from the preceding remarks that in order for a thing *S* to be said to truly have duties it must be the sort of thing which can cognitively access whatever justifying reasons there might be in favor of action. *S* must, in other words, have beliefs or relevantly similar psychological states. This is to say that *S* must be what is called an intentional agent (French, 1984); that is, *S* must be a thing which exhibits intentionality or has intentional states. While it is clearly not the case that all intentional agents have access to justifying reasons—squirrels, for example, seem to be intentional agents—necessarily, if one does have access to justifying reasons for action, one is an intentional agent. Of course, in addition to being intentional agents, in addition to having access to justifying reasons for action, moral agents must also be responsive to or capable of acting on the reasons they have. Hence, if a being is in principle or in fact incapable of acting on the reasons it has (or which we suppose it has), it may be an intentional, though not a moral agent.

In light of the aforementioned remarks, my central argument, which takes the form of a dilemma, may be succinctly stated as follows. (1) Either institutional duty-claims have the same meaning and signification as individual duty-claims or they do not. (2) If they do, then either (a) we must assume that institutions, like individuals, possess and are capable of acting upon reasons for action, or (b) if they are not so capable, we must adopt an error-theoretical perspective according to which all attributions of duties to institutions are meaningful but false. (3) If, however, institutional duty-claims do not have the same meaning as individual duty-claims, then we must admit that such claims involve an equivocation on the term “has a duty.” I will argue that (a) is unsupported—i.e., that we are not justified in assuming it to be true. We are thus left with one of two options: we either admit that institutional duty-claims are false, or we admit that they involve an equivocation. I will also argue that we ought to reject (b) and admit that institutional duty-claims involve an equivocation of terms, which is easily avoided by “translating” them into value-claims.

4. The Case for Institutional Moral Agency

In the previous sections I argued that an individual duty-claim ought to be interpreted as expressing the notion that its subject possesses and is capable of acting upon justifying reasons for action. I went on to explain that this implies that the individual is a rational agent, for only rational agents are able to possess and act upon justifying reasons for action. The question to which I now turn is whether we are justified in supposing that institutions are agents in the sense required for them to be the subjects of duty-claims.

A significant number of scholars have argued that we are so justified. According to those who defend what is variously known as *corporate* or *institutional moral agency* or *moral personhood*, institutional organizations ought to be viewed as moral agents in their own right (see, e.g., Erskine, 2003; French, Nesteruk, and Risser, 1992; Soares, 2003; Wendt, 1999; Wendt 2004).

Most notably, Peter French (1984) has argued that corporations are “full-fledged members of the moral community,” which have “whatever privileges, rights, and duties as are, in the normal course of affairs, accorded to all members of the moral community” (p. 32).

On this view, corporate bodies such as businesses and states have certain goals and they weigh information, and upon this basis they may be said to form intentions to act. The organizational structures and policies that allow for this kind of supposedly deliberative action render corporations teleological or goal-directed systems whose behavior in the world is not properly characterized as a “spontaneous convergence of individual interests” (Erskine, 2003, pp. 23-24; see also, French, 1983). Unlike crowds or mobs, business corporations and governments, for example, may be said to purposively enact policies and expend resources, much in the way individuals purposively seek to satisfy their desires and goals. Insofar as they are voluntary actors who are responsive to reasons, institutions are said to be moral persons.

At the most basic level, French’s argument for institutional moral agency may be understood as unfolding in two steps. First, he argues that corporations are intentional agents, or beings which can form and then act on the basis of intentions. Second, he argues that anything which is an intentional agent is a moral person. Neither of these claims is obvious, so French spends a great deal of time defending each. I will briefly examine the arguments provided for both.

The argument for the first claim—that corporations are intentional agents—may be understood as resting on the assumption that an entity *S* is an intentional agent if and only if (i) *S* has intentional states such as beliefs and desires, and (ii) *S*’s actions are explicated by reference to these intentional states (French, 1984, pp. 39-40). If we wish to explain why an intentional agent did something, we will presumably appeal to her desires and the beliefs she has about how to satisfy those desires. If we are to believe that corporations or institutions are intentional agents, it must be shown that they have intentional states and that it is by reference to these states that we can explicate the institution’s actions. French’s strategy is to identify certain crucial features of institutions which serve as the functional equivalents of what, in human beings, we refer to as beliefs and desires. These functional equivalents are rather obvious, even if they are ultimately inadequate: whereas we have beliefs, institutions create, manage, store, and access *information*; whereas we have desires or ends which we aim to pursue, institutions have goals which are codified in the form of policies or become operative within corporate culture (French, 1984).

On French’s view, the relevant point is that corporations have “corporate internal decision structures” (CID structures), which do two important things: they organize personnel, defining the relationships that exist between offices or departments, and they articulate what is known as “corporate policy” (French, 1984, pp. 41ff). Part of what this involves is setting the rules which govern the exchange of information, recommendations, orders, etc. between the officers or the personnel of a corporation (French, 1984, pp. 42-3). These rules should be understood as primarily creating role responsibilities and duties; they create *individual* responsibilities and duties only in a derivative sense. While it is true that the chief financial officer in a corporation is responsible for risk management, it will only be true that Jane Doe is responsible for risk management if she occupies the role of chief financial officer; if John Roe instead occupies the office, he will have this responsibility.

To say that a CID structure also articulates corporate policy

is to say that it establishes the “recognition rules” according to which a decision or act is judged to have been done for corporate reasons. For a decision or act of an *individual* to be properly described as the *corporation’s* decision or act, French argues it must accord with what is called “the basic belief of the corporation,” or the corporation’s most basic policy (French, 1984, citing Buzby 1962).

By focusing on CID structures, we are to notice at least two things. First, it is the CID structure which incorporates the individual actions of the members of the corporation in such a way that they may be said to be collectively pursuing corporate goals or engaging in corporate projects. It is because a CID structure of a certain sort is in place that individual persons engage in some of the acts they do (i.e., their “official acts”), and it is only because a particular set of rules is in force that individual actions “count” as official acts (e.g., the raising of hands in a particular context counts as voting). Second, the CID structure allegedly allows us to attribute intentionality to a corporation. Suppose, for example, a sufficient number of relevant individuals (e.g., board members) vote in a certain way; this typically counts as the corporation deciding to do something. Now suppose that such a decision is implemented:

[W]hen the corporate act is consistent with an instantiation or an implementation of established corporate policy, then it is proper to describe it as having been done for corporate reasons, as having been caused by a corporate desire coupled with a corporate belief and so, in other words, as corporate intentional. (French, 1984, p. 44)

What French takes to be importantly relevant to the issue of institutional moral agency is that CID structures give rise to information processing and decision-making procedures which are potentially *distinct* from those which isolated individuals would otherwise engage in, and these in turn result in decisions which are not properly attributable to the involved individuals. His point is that insofar as an institution has an internal decision-making procedure, there is a sense in which *it* deliberates. “When operative and properly activated,” he explains, “[a] CID structure accomplishes a subordination and synthesis of the intentions and acts of various biological persons in a corporate decision” (French, 1984, p. 41). In other words, it *incorporates* their individual actions and intentions, and it does so in such a way that the personal interests or parochial concerns a particular employee or executive might have get “diluted” by the subordination process (French, 1984, p. 44). Indeed, it is not difficult to imagine a situation in which the individual members of an institution who participate in decision-making may, as a group and under the influence of institutional roles and rules, reach a decision that no particular individual finds appealing. In such cases, we may say that the institutional decision diverged from the individuals’ own preferences.

We may summarize the key elements of the aforementioned argument in the following way. In subsequent sections, I will refer to this as the argument for intentional agency. As we will see, criticisms of this argument will focus on the second premise.

(1) An institutional organization is an intentional agent if and only if (i) it has intentional states such as beliefs and desires, and (ii) its actions may be explicated by reference to these intentional states.

(2) (i) Institutional organizations have intentional states (institutional *beliefs* take the form of information; institutional *desires* are codified in the form of policies or the operative goals within corporate culture), and (ii) it is by reference to an institution’s beliefs and desires that we explicate an institution’s actions.

(3) Thus, institutional organizations are intentional agents.

French's argument in favor of the notion that anything which is an intentional agent is a moral person is rather complicated and drawn out—something which perhaps should not be surprising given the fact that the claim is rather unintuitive. I say that it is unintuitive because it is commonly supposed that *not* all agents are moral agents. For example, many people believe that most if not all mammals are agents: dogs, cats, beavers, bears, and horses are regularly thought of as agents, since they seem to be the sorts of things which engage in action. In other words, they are all thought to be things which have intentional states (e.g., desires or volitions) which lead them to behave in at least some of the ways they do. This is a point which has been made by Thomas Donaldson (1982):

Some entities appear to behave intentionally which do not qualify as moral agents. A cat may behave intentionally when it crouches for a mouse. We know that it intends to catch the mouse, but we do not credit it with moral agency... One seemingly needs more than the presence of intentions to deduce moral agency. (p. 22; quoted in French, 1984, p. 165)

French himself claims that even if we admit that cats and other animals act in ways that seem to suggest that they can engage in intentional behavior, they are not “full-blooded intentional actors,” since they do not seem capable or free to select their mode of behavior in light of reasons for action (French, 1984, p. 166). It has not gone unnoticed in the literature that his treatment of this issue is underdeveloped and perhaps deeply erroneous (Wall, 2000). However, the task of clearing up this aspect of his argument need not concern us. What must be noted is simply that he takes the view that a thing is an intentional agent only if it is *very* much like a human being in terms of its ability to engage in deliberative behavior that is backed up by reasons. As he puts it, to say a thing is a person is just to say it is a thing whose behavior is explicable by appeal to “a coherent set of true empirical generalizations,” viz., those of folk psychology (French, 1983, p. 249). Once the particular way in which French is using the term “intentional agent” is understood, his assertion that anything which is an intentional agent is a moral person becomes less problematic, since it amounts to little more than the claim that anything which is very much like a normal (adult) human being is a moral person. The significance of this claim arises when we notice how it figures into his overarching argument: in claiming that corporations are intentional agents, French is presumably suggesting that they not only have intentional states such as beliefs and desires (or the functional equivalent of those states), but that they are capable of or free to select a mode of behavior in light of reasons for action.

5. Evaluation of the Argument for Institutional Moral Agency

French's argument for corporate *moral* personhood has received a great deal of critical scrutiny. Since the argument for moral agency hinges on the soundness of the argument for intentional agency, I will focus only on those criticisms that have called into doubt the second premise of the argument for *intentional agency*. To call the second premise into question is to challenge the claim that we are justified in supposing that institutions meet the necessary conditions for intentional agency. Various commentators have done just this, arguing that we have little to no reason to suppose that corporations or institutions really “desire” anything, or “possess reasons,” or “have an interest” in anything.

Let's consider first the issue of institutional goals. French is of the view that corporations may be said to have reasons for action because “they have interests in doing those things that are likely to result in the realization of their established corporate goals...” (French, 1984, p. 45). Moreover, a corporate or institutional “decision” can only be recognized as a corporate decision (as opposed to the decision of mere individuals) if it comports with what is called the “basic belief of the corporation.” This is a term French found in the work of G.C. Buzby, who argued that the interactions, decisions, and goals of individuals involved in a corporate enterprise give rise to or serve to create a corporate “image” that serves as the basis for making judgments concerning whether a putative corporate policy may be attributed to the corporation itself. Buzby suggests that when a more particular policy does not comport with a corporation's basic policy, “it is no longer the policy of that company” (French, 1984, p. 43; citing Buzby, 1962, pp. 5-12). The point is that when a policy or decision made by individuals within the corporation does not comport with the basic goals of the corporation, that policy or decision is more properly attributable to the individual decision-makers and officers, rather than the corporation itself. This view, of course, assumes that corporations do, in fact, have goals or basic policies that *are* properly attributable to the corporation itself. More generally, it assumes that the corporation has a personality which is not immediately reducible to the personalities of its individual member-constituents.

It is here that we ought to take pause. Although French anticipates that his readers may fall under the spell of an “anthropocentric bias” when they reflect on the claim that corporations have goals and desires, he does little to convince us that we are in an epistemic position to grant that corporations and institutions literally have intentional states of this sort which are not reducible to the goals and interests of involved individuals. Michael Keeley (1981) has stated the difficulty facing French's view rather nicely. We must, he points out, distinguish between the goals people have *for* an organization and the notion that there are goals *of* an organization. The goals *for* an organization are goals or preferences possessed by natural people (we might refer to them as the stakeholders). We need not restrict ourselves to shareholders (owners) or employees when we survey the various goals people have for corporations. As Keeley notes, many people in various different positions in a community often have goals or at least expectations concerning any particular institution or corporation. While owners may view a corporation as having the goal of making a return on their investment, employees may view it as having the goal of providing a stable and livable wage. Consumers, of course, tend to describe the goal of a corporation in terms of the services or goods which it makes available in the marketplace (Keeley, 1981, p. 150). The goals *of* an organization would be those that we claim the organization itself (somehow) possesses. French's view requires that there be goals of an organization—i.e., outcomes intended by the organization itself. Keeley raises the skeptical worry that we have no way of determining what these alleged goals are *except* by reference to the goals actual people have for the organization. We can identify the goals for an organization by surveying the various stakeholders and participants in the organization. But, Keeley argues, “it is not apparent that the... goals of an organization... can be identified by any means” (Keeley, 1981, p. 150). Even if we suppose that the goals of an organization can be found in official documents (e.g., charters, annual reports, etc.), we need only recognize that these documents and statements were crafted by individuals and presumably they describe participants' goals for the organization. Or to put the matter in

slightly different terms, we seem wholly justified in supposing that these documents do express or describe some participants' goals, yet we lack any epistemic reason to suppose they express or describe the corporation's own goals. There appears to be a gap between claims about what individuals want and what a corporation wants, and the present point is that this gap cannot be bridged by amassing further evidence about what individuals want or the goals they have.

Keeley imagines that a defender of institutional agency might insist that we can figure out what the organizational procedures are which govern organization behavior and infer from those the "operative" (but perhaps unarticulated) goals of a corporation or institution. From these, we can then derive organizational intentions. But of course, inferring from behavior what the "rules of the game" are (whether they are articulated or not) does not reveal that the organization itself has intentions, nor does it reveal anything by way of what "the game itself intends" (Keeley, 1981, p. 151).

Edmund Wall (2000) has presented a criticism of a similar form, pertaining to the claim that institutions themselves, rather than individuals, make decisions. In response to this claim, Wall points out that we must distinguish between two importantly different claims: on the one hand, we might say that a group of people may arrive at a joint intention by following an agreed-upon process of deliberation; on the other hand, we might say that there is some entity which is not identical to the group of individuals which employs a process of deliberation. A group of individuals may have a decision-making procedure amongst themselves which, in fact, directs their individual actions so that certain results can be achieved, but that does not imply that there is some entity beyond those individuals which actually makes decisions (Wall, 2000, p. 189).

A similar thing can be said about the alleged "beliefs" of an institution, which are said to take the form of the information which is processed within an organization. Although the members of an organization may process and synthesize information, form beliefs, frame conclusions, etc., we need not suppose that some further entity has done any of those things.

Each of the aforementioned criticisms challenges the notion that we have reasons to suppose that institutions possess the cognitive or intentional states requisite for agency. They are compatible with an even broader objection. In a nutshell, the objection is that we have no epistemic reasons to suppose that institutions are conscious as opposed to non-conscious, and since we necessarily imply that a being is conscious when we claim that it is an intentional agent, we ought to be highly suspicious of the claim that the being in fact has the intentional states it is alleged to have.

It is well-known that we owe our notion of intentionality—that is, the *directedness* or *aboutness* of mental phenomena—to the late-19th and early-20th century German philosopher, Franz Brentano. Brentano (1995) famously claimed that desires, judgments, perceptions, emotions and so forth are mental as opposed to physical states insofar as they are intentional states, or "acts of presentation" (pp. 60-61). Whereas physical states cannot really be *about* anything (except when *taken* to be about other things by conscious beings), mental states standardly *are* directed toward or about things, states of affairs, or properties (simple pain and pleasure are obvious examples of mental states that seem to lack this kind of directedness). Such intentional states are inescapably conscious states; there is no sense in which one state can be about another state (or thing or property) without it being about that other thing (etc.) for someone. Now, as I indicated above, we may suppose that *S*

is an intentional agent if and only if (i) *S* has intentional states such as beliefs and desires, and (ii) *S*'s actions are explicated by reference to these intentional states. It seems to me that the possession of an intentional state on *S*'s part entails in some sense that *S* is conscious: *S* cannot be said to have an intentional state unless *S* is a conscious being. If we are unjustified in supposing that a thing is a conscious being, I take it we are thereby unjustified in supposing that it is an agent. This point accords with common sense; we would reject the possibility that trees, for example, are agents not primarily on the grounds that they do not *do* anything—we regularly talk about what trees do—but rather on the grounds that trees do not have beliefs, desires, or other conscious mental states. While we can speak about the states of a tree bringing about certain outcomes, and in that limited sense make true claims about what trees do, we do not take ourselves to be justified in claiming that their doings are caused or properly backed up by mental states.

With this in mind, I wish to suggest that recognition of the fact that we are not in an epistemic position to ascribe to institutions the sort of consciousness which the agency theorist requires can serve as a check on their claims concerning the possession of intentional states by institutions. In that way this "consciousness objection" serves to undermine the force of the arguments provided in defense of institutional moral agency.

The consciousness objection proceeds by making a claim about institutions which is similar in form to the claim made above about trees. While we can speak about the internal decision-making structures that (at least partially) constitute an institution bringing about or making possible certain outcomes, and thus in this limited sense we can make true claims about what institutions do, we are not justified in claiming that the outcomes of individuals acting within an institutional structure are caused by mental states of the institution itself (rather than the mental states of its participants, stakeholders, or other relevant individuals).

The matter is very different when we consider the individuals themselves. We feel confident and justified in claiming that certain of the things which individuals do are caused by their mental states for at least two reasons. First, and perhaps most intuitively, we rely on a kind of analogical reasoning. Each of us believes that certain of our *own* actions are caused by antecedent mental states. For example, I have typed this English sentence because a moment ago I found myself desiring to provide an example of behavior on my part that is caused by one or more conscious mental states—in this case, a desire, as well as the belief that typing this sentence will cause my readers to recognize the sort of phenomena currently under discussion. When I see other people engaging in behavior which is very similar to my own, I justifiably (though fallibly) conclude that their behavior is also caused by conscious mental states. The second reason we are justified in claiming that certain of the things which individuals do are caused by their mental states is related to the first in an important way. In addition to recognizing an analogy between my own case and the case of another, I also recognize that I have no other viable explanation available to account for the other's behavior. When I read my students' essays, for example, the only viable explanation I have available to account for the fact that they have handed me sheets of paper containing markings that I interpret as English words strung together in sentences and paragraphs is that they had a certain understanding of my expectations concerning what they were to do over the weekend as well as certain desires and beliefs concerning how to meet those expectations. Of course, to say that I lack a viable *alternative* explanation for this behavior is not to preclude me

from admitting that there is some *further* explanation or account to be given concerning the underlying causes or states of affairs that make it the case that my students have the mental states I imagine them to have. I need not import any philosophical preconceptions concerning mental substances, mental properties, or the reducibility or irreducibility of mental states to brain states, or so on. And I need not necessarily deny that any such further account may be true.

Returning to the case of the institutions, we find ourselves lacking either of these two bases for ascribing conscious mental states. While we can and do construct narratives in which we anthropomorphize institutions, we do not find ourselves observing the behavior of institutions in a way that would justify drawing an analogy between our own inner mental life and the possible inner mental life of the institution. What we *do* see are fellow human beings acting. This speaks to the second issue: since the only behaviors we observe when we look to see what an institution is doing are the behaviors of our fellow human beings, we have a ready-at-hand explanation for these institutional goings-on. Namely, we can appeal to the same sorts of psychological explanations which we would use to account for our own behavior. What is to be taken note of is the fact that we are not justified on pain of inexplicability to suppose that institutions themselves have an inner mental life. We can explain institutional goings-on by appeal to facts about individuals and their relations. In the case of these individuals, by contrast, we *are* justified on pain of inexplicability in supposing that they each have an inner mental life. We take it that observations of at least some of their behaviors require attributing to them conscious mental states.

Against this line of argumentation, a defender of the institutional agency theory will argue that the “doings” or “decisions” of a corporation are not always immediately attributable to particular individuals, but rather to the institutional structure which is to be taken as the core or essence of the institution. Hence a robust and fully adequate explanation of corporation “action” does place us in the position of needing to posit that the institution possesses intentional states, else these institutional outcomes would be unexplained. Let us carefully examine how this story is supposed to go to see whether this conclusion is inescapable.

The first task is to explain the sense in which certain doings or decisions are not properly attributable to individuals. As we saw, French correctly supposes that when individuals find themselves within an institutional structure, their personal values, desires, and beliefs can be effectively silenced or rendered inoperative through the incorporating efforts brought about by the policies and rules which govern their official activities. For example, it is not uncommon for academics to accept positions at private colleges which are associated with religious organizations to which the academic herself does not belong. While the individual may not share the values or beliefs which the institution as a whole is expected to promote, it often remains possible for her adequately to fulfill the role of, say, a philosophy or history professor in the institution, and even to contribute in a role-specific way to the promotion of the institution’s religious goals (e.g., by offering certain courses relating to the religious worldview promoted at the institution). Similarly, those who are familiar with Bernard Williams’s (1973) criticisms of utilitarianism will be familiar with his hypothetical situation concerning a chemist, George, who is morally opposed to the use of biological weapons, and yet finds himself presented with the opportunity to take up a position in a weapons manufacturing facility. Williams fully accepts that George could, conceivably,

fulfill the role-specific duties associated with such a position despite his personal commitments (Smart and Williams, 1973, pp. 97-98). Although situations like this may raise interesting problems concerning moral integrity—a point which Williams makes—the following point seems largely beyond doubt: as long as a role is defined by rules that govern behavior and activity, not beliefs or values, then it is largely unimportant which particular beliefs or values an individual occupying the role happens to have. Such an individual can usually be caused to act in ways that she otherwise would not have were she not placed in the institution in precisely that way. If an institution is understood primarily as a set of offices or roles, and the behavior of those who fill these offices or roles is governed in this manner, it is both conceptually and practically possible for individual officers within the institution to make decisions that diverge from the decisions they would have otherwise made.

Let’s carefully note what is being said here. The foregoing considerations simply reveal that it can end up being the case that a group of individuals, standing in the relations made possible by the rules and policies under which they operate, will, individually or in tandem, reach decisions and bring about outcomes that they would not have reached or brought about were it not for the fact that these rules and policies were in place. The defender of institutional agency wishes to suggest that this evidences the fact that some of the decisions and actions of individual corporate actors are thus more properly understood as being caused by the institutional arrangement, rather than the individuals themselves. We need not and should not deny any of this; institutional arrangements clearly *do* play a causal role in influencing individuals’ actions. That having been said, I do not see how an appeal to the joint activity of individuals or their responsiveness to institutional circumstances and demands justifies us in making attributions of even the most rudimentary or lowest forms of consciousness or intentional states to institutions. As long as institutions are populated by conscious, intentional actors, explanations of corporate outcomes are forthcoming which make reference to their beliefs, desires, and so forth.

The agency theorist’s case may, at least on the face of it, be strengthened by considering a thought-experiment. If it can be shown that it is conceivable that an institution could continue to operate without relying on individuals at all, that might seem like persuasive evidence in favor of attributing intentional states to that institution. Patricia Werhane (1985) has offered a fictional scenario that may be put to precisely this end. (Note, however, that this is not the purpose to which Werhane herself puts this fictional account. Like me, she is engaged in criticism of French’s view, and she uses this scenario to conclude that institutions are not free agents.)

One might imagine a corporation that was operated solely by robots and computers. Such an organization, let us call it Robotron, would have a charter and legal status. It would operate like other corporations. It would own property, manufacture products, conduct marketing, correspond with other corporations and with customers, replace obsolete equipment, develop new product lines, write proxy statements, answer SEC inquiries, etc. Robotron would have stockholders and pay out dividends. It could draw up rules for robot-corporate behavior, and could develop corporate goals and a hierarchy. The rules and structure of its electronic decision-making could be such that one might call Robotron an intentional system. (p. 38)

It would clearly be inappropriate to claim that statements concerning the beliefs, desires, and actions of Robotron are reducible to statements concerning the present members of Robotron, and this is due to the simple fact that Robotron has no

members which are individual human beings. Moreover, in this case, like in the cases of other institutions, it would be inappropriate to conclude that the institutions' "doings" or decisions are random. So, does this thought-experiment thereby count as a case in which it would be appropriate to ascribe to the institution itself—to Robotron—conscious mental states? Should we conclude that Robotron has intentional states? I think not. Indeed, I think this thought-experiment may help us see precisely what is at issue in this controversy.

Robotron clearly does things, at least in one sense of that phrase. Things happen in the world as a result of Robotron remaining in operation; goods are produced and released in the market, money is deposited into shareholders' accounts, and so on. The appropriate and clearly justified way of stating what is going on here is that Robotron's organizational structure, the input of resources, and so forth cause certain outputs, and there is clearly a causal relationship between the outputs and Robotron (that is, the organizational structure, the policies, rules, and goals that largely constitute Robotron). Yet notice that this merely establishes that corporate outputs may be given a mechanistic explanation, and such a mechanistic explanation does not imply agency. Here again we find ourselves utterly without grounds upon which to believe that Robotron has conscious mental states, for the mechanistic explanation is a wholly adequate explanation. There is no explanatory gap that must be filled by positing that Robotron has internal mental states.

To head off a possible objection, let me note that none of this should be construed as denying the potential utility of using intentional language to describe and predict institutional outcomes. As a way of defending the notion that institutions are intentional agents it is sometimes argued that the language of collective intentionality has predictive power. It is well known that Daniel Dennett (1989) has advocated adopting an "intentional stance" when we are trying to predict future behavior of a complex system. Dennett explains what is involved in adopting this stance:

first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in most instances yield a decision about what the agent ought to do; that is what you predict the agent will do. (p. 17)

Dennett is careful to note that it is not the case that all systems whose behavior is susceptible to being cast in intentional language in fact have intentional states. However, it has become somewhat popular for defenders of institutional intentional agency to rely on the predictive power of the intentional stance when making their case. Deborah Tollefsen (2002), for example, uses this "interpretationalist" approach to defend the notion that corporations are "true believers." But such an approach takes too narrow of a view. What is at stake in the debate over intentional agency is not whether it is sometimes or even oftentimes useful to describe institutional behavior using the language of intentionality. It is, rather, whether we are ultimately justified in supposing that the attributions of intentional states ought to be construed literally. The availability of mechanistic or quasi-mechanistic explanations speaks strongly against us being so justified. When we contemplate the doings of institutional organizations and when we contemplate the doings of our fellow human beings we do not find ourselves in precisely the same epistemic position. In the case of the former,

we are not justified on pain of inexplicability in supposing that they have an inner mental life.

Let us now return to the "straightforward" interpretation of institutional duty-claims according to which they have the same meaning and signification as individual duty-claims. If my argument is sound, we must notice that this interpretation requires that we affirm precisely the view which I have just rejected: in order for attributions of duty-claims to turn out to be (literally) true, it must be the case that institutions have mental lives. The aforementioned objections to the argument for intentional agency constitute the basis upon which we can formulate a *reductio* argument against this view. This forces us into the position of having to adopt one of two alternative interpretations. The first is what I referred to at the outset as the error-theory: we could *admit* that the "straightforward" interpretation captures the *meaning* of institutional duty-claims, but given the implausibility of the argument for intentional, we could *also* admit that all such institutional duty-claims will turn out to be (literally) false. The second option is to entertain the possibility that when we say that institutions have duties we *do not* mean anything so implausible; putative institutional duty-claims are sometimes true and yet they do not have the same meaning and signification as individual duty-claims. This second alternative strikes me as the more appropriate of the two. I will thus conclude by briefly sketching an interpretative framework of this sort.

6. An Alternative Interpretation

Whereas the analysis of institutional duty-claims which I have been criticizing only *seemed* straightforward and commonsensical, the alternative that I wish to propose is *genuinely* straightforward and commonsensical. It is simply this: when someone says of an institutional organization that it has a duty to do *a* or bring about *g*, they simply mean that a possible organizational structure which will result in *a* being performed or *g* being pursued would be *good*; that is, it would be productive of value. For example, when one says that BP has a duty to compensate the victims of the Gulf Oil Spill, this may be interpreted as expressing the claim that BP being organized in a way that would allow for this outcome to be achieved would be valuable or is desirable for moral reasons.

Of course, such a claim typically carries with it the implicit suggestion that not only would it be good, but that it thereby ought to be brought about. To continue with the example, when it is said that BP has a duty to compensate victims of the oil spill, we should understand ourselves to be saying that *we* (or some other relevant party or parties) ought to make it the case that BP carries out such an action or pursues such a goal. When a putative institutional duty-claim is used in this way it is not really about the duties of the institution (for, as we have seen, we lack good reasons for supposing that institutions are the sorts of things that can have duties); rather, it is a claim about the kinds of institutional structures that we (or other relevant parties) are morally justified in putting into place. In other words, the statement, although it appears to be about the institution's reasons for action, is actually better understood to express a value-claim—that is, a claim about the utility or value of modifying the institutional structure—which serves as a basis for a claim about the reasons people, working either individually or in tandem, have for acting. Such assertions express the reasons we have for crafting institutional policies and conferring upon individuals who occupy offices within our institutions new and perhaps hitherto unusual duties, obligations,

and rights. The plausibility of this interpretive framework rests upon the recognition that it captures everything which we want and need to say when we engage in ethical discourse concerning institutional organizations.

References

- Brentano, F.C. (1995), *Psychology from an Empirical Standpoint*, Routledge, New York.
- Buzby, G.C. (1962). "Policies—A Guide to What A Company Stands For." *Management Record*, Vol. 24, pp. 5–12.
- Dennett, D.C. (1989), *The Intentional Stance*, MIT Press, Cambridge.
- Donaldson, T. (1982), *Corporations and Morality*, Prentice-Hall, Englewood Cliffs.
- Erskine, T. (2003) "Assigning Responsibilities to Institutional Moral Agents: The Case of States and "Quasi-States", in Erskine, T. (Ed.) *Can Institutions Have Responsibilities? Collective Moral Agency and International Relations*, Palgrave Macmillan, New York, pp. 19–40.
- Fink, L. (2012). "Companies Have a Moral Duty to Help Workers Save." *Financial Times*, 28 Feb. Available <http://www.ft.com/intl/cms/s/29de9e58-61fd-11e1-807f-00144feabdc0,Authorised=false.html>.
- French, P.A. (1983). "Kinds and Persons." *Philosophy and Phenomenological Research*, Vol. 44, pp. 241–254.
- French, P.A. (1984), *Collective and Corporate Responsibility*, Columbia University Press, New York.
- French, P.A. (1995), *Corporate Ethics*, Harcourt Brace, Fort Worth.
- French, P.A., Nesteruk, J., Risser, D.T. (1992), *Corporations in the Moral Community*, Harcourt Brace Javanovich College Publishers, Fort Worth.
- Keeley, M. (1981). "Organizations as Non-Persons." *Journal of Value Inquiry*, Vol. 15, pp. 149-155.
- Ross, W.D. (2002), *The Right and the Good*, Oxford University Press, Oxford.
- Sherlock, M.F., Lunder, E.K., Liu, E.C., and Klein, H.S. (2010). "Tax Issues and the Gulf of Mexico Oil Spill: Legal Analysis of Payments and Tax Relief Policy Options." Congressional Research Service. Available <http://www.fas.org/sgp/crs/misc/R41323.pdf>.
- Smart, J.J.C., Williams, B.A.O (1973), *Utilitarianism. For and Against*, Cambridge University Press. Cambridge.
- Soares, C. (2003). "Corporate Versus Individual Moral Responsibility." *Journal of Business Ethics*, Vol. 46, pp. 143-150.
- Tollefsen, D. (2002). "Organizations as True Believers." *Journal of Social Philosophy*, Vol. 33, pp. 395–410.
- Wall, E. (2000). "The Problem of Group Agency." *Philosophical Forum*, Vol. 31, pp. 187–197.
- Wendt, A. (1999), *Social Theory of International Politics*, Cambridge University Press, Cambridge.
- Wendt, A. (2004). "The State as Person in International Theory." *Review of International Studies*, Vol. 30, pp. 289–316.
- Werhane, P.H. (1985), *Persons, Rights, and Corporations*. Prentice-Hall, Englewood Cliffs.

Author

Dr. Christoffer S. Lammer-Heindel, Instructor of Philosophy, Loras College, 1450 Alta Vista Street, Dubuque, Iowa 52001, USA. E-mail: christoffer.lammer-heindel@loras.edu. Dr. Lammer-Heindel received his PhD in Philosophy from the University of Iowa in 2012. His primary research interests center on matters of ethics and social-political philosophy. He has taught a variety of applied ethics courses, and he is currently teaching critical reasoning and contemporary philosophy.