Why non-human agency?

3 Why Non-Human Agency

A. Waltermann

1 Introduction

There are increasingly more entities created by human beings but not themselves human that shape our lives, our social and legal reality, such as tables, books, computers, corporations, states and artificial intelligence. Next to churches, states and corporations have been on this list for centuries, but stock-trading algorithms, autonomous weapons and self-driving cars are newer additions. From a legal perspective, churches, states and corporations as well as artificially intelligent entities have in common that they raise the question whether and how such entities are and should be regulated. This entails questions such as whether and how they can act and when they are and should be liable for their actions. From a moral perspective, one can ask similar questions: can these entities act and are they – or should they be – morally responsible for their actions?

Questions of moral responsibility or legal liability are on the public, political and academic agenda, with the European Parliament calling for 'electronic personhood'[1] for autonomous and artificially intelligent robots and with self-driving cars beginning to drive on US roads, but also with an on-going debate regarding corporate criminal liability in Germany,[2] to give only three examples. In the context of moral responsibility or legal liability, it is often (but not necessarily) presupposed that the entity in question has acted, or more generally that there was conduct of some kind. This is not obvious, however. Are non-human entities able to act for the purpose of moral responsibility or of legal liability?

In this paper, I want to focus on the topic of acts of non-human entities and sketch out why I consider this a relevant field of research for academics from law, philosophy, the social and the cognitive sciences in addition to the questions of responsibility that are already being asked. As such, this paper is meant to identify relevant research questions and to out a research agenda. It is not yet meant to answer these questions, only to provide impetus for future research.

1 European Parliament, Report with recommendations to the commission on Civil Law Rules on Robotics, A8-005/2017

2 *Cf.* K. Volk, K. LuDerssen & Eberhard Kempf, *Unternehmensstrafrecht*, De Gruyter, Berlin 2012.

I seek to identify areas in which different fields can learn from one another, in particular areas in which law can learn from philosophy and the cognitive sciences broadly construed, and to identify questions that are relevant to ask in connection to non-human agency. In order to do so, I will first look at examples of different types of non-human entities that have historically or are currently regarded as agents in the eye of the law (in different legal systems), examining the reasons why these have been or are regarded as such (Section 2). This examination reveals a number of (metaphysical) assumptions commonly made regarding agency (both human and non-human), forming the 'metaphysical frame of reference' of agency (Section 3 and 4). This paper will posit (in Section 5) that agency in the law is constructed, but that its construction is often based on an unqualified and uncritical application of the metaphysical frame of reference and is as such insufficiently theorised. Insights from philosophy and the cognitive sciences broadly construed are relevant for an investigation of the metaphysical frame of reference and in future theorising on the matter. The paper concludes in Section 6 by summarising the main argument and pointing out questions for future research.

2 Non-Human Actors and the Law

In everyday life, we routinely make a distinction between agents and things. This holds intuitive appeal. Greene and Cohen summarise:

> Looking out at the world, it appears to contain two fundamentally different kinds of entity. On the one hand, there are ordinary objects that appear to obey the ordinary laws of physics: things like rocks and puddles of water and blocks of wood. These things do not get up and move around on their own. They are, in a word, inanimate. On the other hand, there are things that seem to operate by some kind of magic. Humans and other animals, so long as they are alive, can move about at will, in apparent defiance of the physical laws that govern ordinary matter. Because things like rocks and puddles, on the one hand, and mice and humans, on the other, behave in such radically different ways, it makes sense, from an evolutionary perspective, that creatures would evolve separate cognitive systems for processing information about each of these classes of objects. There is a good deal of evidence to suggest that this is precisely how our minds work.[3]

3 J. Greene & J. Cohen, 'For the Law, Neuroscience Changes Nothing and Everything', *Philosophical Transactions: Biological Sciences*, 2004, pp. 1775-1785, at p. 1782, references omitted.

Chopra and White similarly hold that "[i]ntuitively, an agent is something able to take actions. One way to distinguish agents from other entities is that agents do things, as opposed to have things happen to them."[4] This distinction between agents and other entities seems to correspond to the animate-inanimate distinction outlined by Greene and Cohen.

When it comes to agency in law, I want to distinguish between two notions of agency: on the one hand, there is the notion of agency that relates simply to doing things in the eyes of the law, as a baseline conception of agency; on the other hand, there is a more specific notion of agency used more frequently in the legal context: that of agency in the sense of an agent – principal relationship in the law of agency. These two uses of 'agency' are not always clearly distinguished.[5] By the baseline conception of agency, I mean that the entity is regarded as "a being with the capacity to act",[6] and by act I mean that there can be conduct by that entity.[7] Being an agent in an agent – principal relationship presupposes this more basic sense of agency: here, the agent acts on behalf of another, the principal.[8] In this paper, I am not interested primarily in the more specific notion of agency in the sense of an agent – principal relationship so much as in the more fundamental question why the law regards an entity to be capable of action at all.

Currently and over the centuries, different sets of entities have been regarded as legal agents. It is not the case that all human beings have been or are regarded as legal agents at all times, nor is it the case that only human beings have been or are regarded as legal agents. In the following, I will consider three non-human entities that are or have been regarded as legal agents: corporations and States as current examples and animals in the Middle Ages as a historical example.

2.1 Corporations

It is no exaggeration to point out that corporations are ubiquitous in modern society, or that they impact human lives and the environment more generally on a (near-)constant

4 S. Chopra & L. White, *A Legal Theory for Autonomous Artificial Agents*, University of Michigan Press, Ann Arbor 2011, p. 11.

5 By way of example, *see* id. Perhaps they did not see the need to make the distinction.

6 M. Schlosser, 'Agency', *in* E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2015 edn.; https://plato.stanford.edu/archives/fall2015/entries/agency/.

7 Conduct includes omissions. I will ignore the complication of omissions for the rest of this paper.

8 For an amusing recapitulation of agency theory, *see* S.P. Shapiro, 'Agency Theory', *Annual Review of Sociology*, Vol. 31, 2005, pp. 263-284. *Cf. also* https://en.wikipedia.org/wiki/Law_of_agency, last accessed 1 March 2019.

basis at this point in time. The fields of corporate law (in particular corporate criminal law) and corporate moral responsibility have considered the question whether corporations can act and offer different conceptions of the corporation that lead to different answers.

Criminal law has traditionally been perceived of as ''designed for natural persons, as only individuals can act, so it is argued, and only human beings are considered to be morally blameworthy''[9] but has since moved on to recognize corporate crime in many jurisdictions. As such, the way corporate criminal law conceptualizes corporations and corporate wrongdoing (which, notably, includes *doing* in some form) offers fertile grounds for investigating the reasons why law – at least this field of law – regards corporations as agents, or why it does not.

There are two main conceptions of the corporation here: the nominalist approach and the organisational approach.

The nominalist (or atomist) approach regards the legal entity as nothing more than a legal fiction. Under this view, the corporation is nothing more than a sum of its human parts, a collection of individuals instead of an entity separate from the individuals that make up the corporation. For corporate criminal law, this approach means that the criminal liability of the legal person is 'necessarily derived from the wrongdoing of a natural person'.[10] The *actus reus* (act or omission) and *mens rea* of a natural person are attributed to the legal person, which means that there can be no corporate act (or omission) without a natural person having acted (or omitted to act).

The organisational (or realist) approach views the corporation as a separate entity instead, capable of acting independently from its individual human parts. For criminal liability, this means that the corporation can act and be at fault separately from individual human action and fault; it is not required that an individual's *mens rea* and *actus reus* are first identified and then attributed to the corporation. This approach takes into account the insight that corporations and corporate attitude may influence the conduct of their members.

French considers that to reduce the corporation to ''biological referents'', that is, to human beings is fallacious and argues that the corporation is a non-eliminatable subject of

9 D. Roef, 'Corporate Criminal Liability', *in* J. Keiler & D. Roef (Eds.), *Comparative Concepts of Criminal Law*, 3rd edn, Intersentia, Antwerp 2019, p. 335.

10 Id., p. 336.

responsibility in a way that, for example, a mob is not.[11] This means that while a mob's actions are reducible to the actions of its individual human parts, the same is not true for a corporation, because the mob has – according to French – no intentionality beyond that of its individual human parts, while the corporation does:

> We can describe many events in terms of certain physical movements of human beings and we also can sometimes describe those events as done for reasons by those human beings, but further we can sometimes describe those events as corporate and still further as done for corporate reasons that are qualitatively different from whatever personal reasons, if any, component members may have for doing what they do. Corporate agency resides in the possibility of CID Structure [corporate internal decision structure] licensed redescription of events as corporate intentional.[12]

Because of a corporation's structure of decision making, consisting of an organisational flow chart of corporate power structures and corporate policy, a corporation can be said to act in ways that are different and for reasons that are different than the acts and reasons of the human members. This argument rests on an understanding of actions as those events that can be described under an aspect that makes them intentional. For French, this means that a corporation acts intentionally when its act is ''consistent with, an instantiation or an implementation of established corporate policy'', because it is then proper to describe that act has having been done for corporate reasons, rather than for whatever reasons, if any, the directors or shareholders may have for doing the same thing.[13]

With this, French rejects as unhelpful (for his purposes) a theory that recognises corporations as pre-legal sociological persons:

> Underlying the theory is the view that law cannot create its subjects, it only determines which societal facts are in conformity with its requirements. At most, law endorses the pre-legal existence of persons for its own purposes.[14]

Neumann offers a compelling argument against this approach: because it is a question of decision whether and what legal consequences are attached to a particular social fact or situation, the existence of a pre-legal act by a pre-legal person says nothing about whether

11 P.A. French, 'The Corporation as a Moral Person', *American Philosophical Quarterly*, Vol. 16, No. 3, 1979, pp. 207-15, at p. 210.
12 Id., p. 212.
13 Id., p. 213.
14 Id., p. 209.

this is to be recognised as a legal act or the entity as a legal person.[15] For this, a normative argument is required. Kelsen recognised this when he held that

> To define the physical (natural) person as a human being is incorrect, because man and person are not only two different concepts but also the result of two entirely different kinds of consideration. Man is a concept of biology and physiology, in short, of the natural sciences. Person is a concept of jurisprudence, of the analysis of legal norms.[16]

This can be expanded on for present purposes: in the same way in which 'person' is a concept of jurisprudence, 'action' (or more generally, 'conduct') is also a legal concept, with the law determining what legal consequences attach to an event or situation and whether that event or situation is to be classified as an act (or more generally, as 'conduct'). According to this view, which Neumann calls normativism (*Normativismus*),[17] it is a question of decision whether something counts as a legal act and whether an entity counts as a legal agent. Therefore, the existence of a pre-legal act or agent says nothing about whether this is to be recognised. From a legal perspective, so Neumann, an action is not a natural, but an institutional fact which is constituted by legal or social recognition.[18]

2.2 States

The Draft articles on Responsibility of States for Internationally Wrongful Acts (hereafter: ARSIWA) and commentaries thereto adopted by the International Law Commission in 2001 provide a good starting point to consider the question when a State has acted. They 'seek to formulate, by way of codification and progressive development, the basic rues of international law concerning the responsibility of States for their internationally wrongful acts.'[19] One of the issues covered by these articles includes 'determining in what circumstances conduct is to be attributed to the State.'[20]

15 Volk, LuDerssen & Kempf 2012, p. 16.
16 H. Kelsen & A. Wedberg, *General theory of law and state*, Harvard University Press, Cambridge Massachusetts 1945, p. 94.
17 Volk, LuDerssen & Kempf 2012, p. 16.
18 Id., p. 17.
19 International Law Commission, Draft articles on Responsibility of states for Internationally Wrongful Acts, with commentaries, 2001, p. 31.
20 Id.

Article 1 of ARSIWA states that ''[e]very internationally wrongful act of a State entails the international responsibility of that State.'' The commentary to this article holds that one or more actions or omissions or both may constitute an act in the sense of this article.

Article 2 of ARSIWA states that ''[t]here is an internationally wrongful act of a State when conduct consisting of an act or omission: (a) is attributable to the State under international law; and (b) constitutes a breach of an international obligation of the State.'' The commentary to this article is of particular interest for present purposes. Under (5), it is held that

> [f]or particular conduct to be characterized as an internationally wrongful act, it must first be attributable to the State. The State is a real organized entity, a legal person with full authority to act under international law. But to recognize this is not to deny the elementary fact that the State cannot act of itself. An 'act of the State' must involve some action or omission by a human being or group: 'States can act only by and through their agents and representatives.[21]

According to this, a State is real and can act, but not of itself. It can only act where a human being or group has acted (or omitted to act). This is reminiscent of the nominalist approach to the corporation; in international relations literature, it is called the functional theory and can be contrasted to the agential theory. Fleming summarises the two as follows:

> According to the *agential theory*, states can be held responsible because they are moral agents like human beings, with similar capacities for deliberation and intentional action. State responsibility is essentially individual responsibility scaled up. According to the *functional theory*, states can be held responsible because they are legal persons that act vicariously through individuals. States are principals rather than agents, and state responsibility is akin to the vicarious responsibility of an employer for the actions of her employee. The two theories of state responsibility depend on different understandings of how corporate entities can act.[22]

These two theories of state action and state responsibility closely resemble the realist and nominalist approaches to the corporation.

---

21 Id., p. 35, citing the German Settlers in Poland case, Advisory Opinion, 1923 PCIJ (Ser. B) No. 6, at 22.

22 S. Fleming, 'Moral Agents and Legal Persons: The Ethics and the Law of State Responsibility', *International Theory*, Vol. 9, No. 3, 2017, pp. 466-489, at p. 467, references omitted.

2.3 Animals

In the Middle Ages and the Renaissance, animals were regarded as criminally liable by both secular and religious authorities and prosecuted accordingly, with wild animals generally falling under the jurisdiction of ecclesiastical courts and domestic animals under the jurisdiction of ordinary criminal courts.[23] More than two hundred cases of animal trials are recorded in Western Europe; most from the period of the 15th to 17th century, but with medieval record keeping notoriously spotty, the actual numbers are likely higher.[24] There are documented trials against 'an unspecified number of rats, which were prosecuted in the ecclesiastical court of Autun for having feloniously eaten and wantonly destroyed local barley,'[25] sows and other animals for infanticide,[26] bees, bulls, horses and snakes for homicide,[27] field-mice for fraud and foxes for theft.[28] Punishments for those found guilty ranged from a reprimanding knock on the head to excommunication or capital punishment.[29] A sow, for example, was tried in a court of law for infanticide in Falaise (France) in 1386 and sentenced to death. The sow 'was dressed in human clothes, mutilated in the head and hind legs, and executed in the public square by an official hangman' after '[h]aving been duly tried in a court of law'.[30] Despite the examples mentioned here, it is not the case that these trials were privileging the human plaintiffs or finding for them invariably.[31] These trials, it seems, were conducted in accordance with due process at the time:

> The animals were indeed fortunate, since they received due process of law every bit as good as that provided for humans. They were certainly not executed summarily, as would be the case today of an animal that ate a baby. In fact, the prosecutors were so careful to apply the law equally to animals, that they even placed them on the rack and tortured them for a confession. […] [Animals] were only treated cruelly by the criminal law to the same extent that humans were.[32]

23 W. Ewald, 'Comparative Jurisprudence (I): What Was It Like to Try a Rat?', *University of Pennsylvania Law Review*, Vol. 143, No. 6, 1995, pp. 1889-2149, at p. 1903 *et seq*.
24 Id., p. 1903.
25 P. Beirnes, 'The Law Is an Ass: Reading E.P. Evans' the Medieval Prosecution and Capital Punishment of Animals', *Society and Animals*, Vol. 2, No. 1, 1994, pp. 27-46, at p. 30.
26 Id., p. 31.
27 Id.
28 Id.
29 Id.
30 Id.
31 Ewald 1995, p. 1902.
32 Beirnes 1994, p. 39, citing G. Newman, *The punishment response*, J.B. Lippincott, Philadelpia 1978, p. 93.

This was because ''animals were apparently regarded as having the same sort of responsibility for their actions as humans'', meaning that ''justice and fairness demanded that animals should be treated in the same way as humans.''[33] In sentencing, too, animals were treated similarly to human beings:

> Animals condemned to death were executed in various ways. Some were burnt at the stake; others merely singed and then strangled before the body was burned. Frequently the animal was buried alive. A dog in Austria was placed in prison for a year; at the end of the seventeenth century a he-goat in Russia was banished to Siberia. Pigs convicted of murder were frequently imprisoned before being executed; they were held in the same prison, and under substantially the same conditions, as human criminals.[34]

It seems that animal trials took place from at least the 12th to the 19th century in a variety of Western countries. There are indications, however, of animals being held criminally liable in non-Western countries and long before the Middle Ages: Beirnes cites evidence of animal trials in classical Athens and indicates that it is known that ''the ancient Persians, for example, considered animals as responsible beings, and punished them for their misdeeds.''[35]

It bears keeping in mind that these animal trials took place in a different context from trials today. Cohen (1986) points out that 'one cannot settle in purely empirical terms the question of whether animal trials existed either outside the occident or in pre-medieval Europe' because any such determination depends to a large extent on what is seen as counting as an animal trial.[36]

With this caveat in place, there are some indications why animals were put on trial and prosecuted. Beirnes notes that there is ''no solid evidence of a general belief that the volition and intent of animals was of the same order as those of humans''[37] though apparently, they were believed to be responsible for their own actions in the same way as human beings.[38] The *lex talionis* is one possible – and likely – factor contributing to the trials of animals in the Middle Ages. Similarly,

33 Id., p. 38.
34 Ewald 1995, p. 1905.
35 Beirnes 1994, p. 37.
36 Id., citing E. Cohen, 'Law, folklore and animal lore', *Past and Present*, No. 110, 1986, pp. 6-37.
37 Id., p. 29.
38 Id., p. 38.

> [i]n the Athenian Prytaneum, in the common hearth of the city, a special court was convened to try unknown murderers, inanimate objects such as stones and beams, and animals that had caused human death. In cases of wrongful death it was necessary formally to try these three categories of malefactor because, in reasoning similar to that underlying the *lex talionis*, an unavenged murder would disturb the moral equilibrium and the physical health of the community, the wrath of the Furies would be aroused, and the soul of the deceased would be unable to find rest.[39]

Beirnes further points out that "numerous purposes have been allotted animal trials, quite apart from the use of biblical precepts as justifications for them. Thus, it is said that animal trials were conducted in order to deter other animals from committing similar crimes",[40] but he this considers unlikely, as medieval folk apparently did not "parade their animals in front of the gibbet on which executed pigs were left by the hangman for public display"[41] – or that the animal trials were designed to deter those "responsible for an offending animal's dangerous actions."[42] Ewald rejects explanations that sever the link between an animal's guilt and the animal's punishment as unsatisfactory.[43] When it comes to deterrence as a reason for putting animals on trial and punishing them, he holds that it explains the punishment, but not the trial itself – after all, the animal could have been killed and its body displayed without a trial as well. Animal trials seem to suggest that animals (from cows and wolves to insects and mice) were seen as "creatures who, like humans, could be brought to trial for their deeds and cruelly punished; but from some points of view this must be seen as a sign of moral respect" and as "fellow creatures who enjoyed certain basic rights that can be vindicated at law."[44] Today, we do not think of animals in the same way anymore:

---

39 Id., p. 37.

40 Id., p. 38.

41 Id.

42 As a footnote, it is interesting to point out here that the determination of responsibility – certainly in the case of legal liability – is something a trial would have to establish in the first place. If it is the animal that is capable of being responsible for its own actions, legally speaking or in some other meaning of the word, this phrasing would point back to the animal as the one responsible for its own behaviour. Out of the context, it is apparent that Beirnes here means human beings – in particular the owners of the animal. This is apparent because Beirnes continues: "If so, animal trials served as vehicles to convey to animal owners a moral message to oversee their charges. In a case of 1567 where a pig was executed for having killed an infant in Senlis in France, therefore, the judge warned the inhabitants of the village "not to permit the like to go unguarded on pain of an arbitrary fine and of corporal chastisement in default of payment". *See* id. While context makes clear what the statement refers to, the ambiguity of it hinging on the word 'responsibility' is noteworthy.

43 Ewald 1995, p. 1906.

44 Id., p. 1915.

> […] we draw a sharper distinction between the animals and ourselves, and are more inclined to view them as automata, as parts of the material world. And when we do accord them some degree of moral respect, there has been an important change in the standard we apply: the higher animals are not to be mistreated, not because they are the handiwork of God, but because they are like us.[45]

What lies at the heart of the animal trials and modern difficulties with grasping their rationale full is a difference in what may be termed the metaphysical frame of reference:[46] who we regard as legal agents today and what forms of punishment we are willing to inflict on them seems influenced, to a large extent, by modern moral sensibilities and intuitions. Questioning the rationale behind animal trials and seeking to understand why animals were regarded as legal actors and both tried and punished accordingly invites questioning the rationale of our current metaphysical frame of reference as well.

What is our current metaphysical frame of reference and how does it impact our (legal) view of non-human entities as actors?

3 The Metaphysical Frame of Reference of Action

The law can, in principle, define its own concepts and create its own frame of reference entirely without the need for metaphysics or to refer to an existing frame of reference. This is the point Neumann makes with regard to corporate criminal liability[47] and which Brozek makes regarding personhood in the law when he holds that it "has a technical character and is not connected with any concrete philosophical content."[48] With regard to personhood as well, however, Naffine points out that the law does not do this cleanly or neatly.[49] Instead, it oscillates between different approaches to or models of personhood[50] with different metaphysical "baggage".[51]

45 Id.
46 Id., pp. 1916-1921.
47 Volk, LuDerssen & Kempf 2012, p. 16 *et seq*.
48 B. Brozek, 'The Troublesome Person', *in* V. Kurki & T. Pietrzykowski (Eds.), *Legal Personhood: Animals, Artificial Intelligence and the Unborn*, Springer, New York 2017, p. 8.
49 N. Naffine, 'Who Are Law's Persons? From Cheshire Cats to Responsible Subjects', *The Modern Law Review*, Vol. 66, No. 3, 2003, pp. 346-367.
50 Id.
51 Brozek 2017, p. 6.

While the law may assign an entity the status of agent independently of concrete philosophical content or particular physical characteristics, this can (and often will) be influenced by philosophical thought and folk psychological understandings of the notion of agency on the one hand; it also can (and often will) influence philosophical and folk psychological understandings on the other hand. While the law is independent of philosophy, moral theory, folk psychology, etc. this does not take away the possibility (and indeed likelihood) of mutual influence.[52]

We have seen above that when it comes to the agency of corporations and States, the law shifts between different models as well. In corporate law, particularly corporate criminal law, the nominal approach seems to assume that only events caused by human beings can count as actions proper and that only human beings can have intentions. It accepts that those actions can then be attributed to a composite entity such as the corporation of the state and in this sense, corporations and states can and do act, but their acts are always derivations of and reducible to human acts. In contrast, the organisational approach does not take human beings as the prototypical agents on whose actions all other forms of action must be based. International law is consistent in treating States as incapable of non-derived actions. In international law, a State can only act where its acts are reducible to human acts or, conversely stated, where human action is attributable to the State. This could be taken as emphasising the distinction between the legal notion of action and other conceptions, but this strikes me as the wrong conclusion to draw. The phrasing of the International Law Commission of it being an ''elementary fact'' that an act of a State must be preceded by human action that is attributable (and actually attributed) to the State suggests not a deliberate choice for distinctness from other fields. Instead, it seems a reference to a metaphysical frame of reference that is presumed to be ''fact'' – and elementary fact at that.

Synthesising the findings from the sections on corporations, states and animals above, two approaches can be identified with regard to legal agency. Under both approaches, non-human entities can and do act. The difference lies in how they are able to perform an action, that is, in the conceptualisation of their action. Under one model, non-human entities act only when the act of a human being is attributed to them. The action of a human being counts as the action of a corporation or a state and as such *is* the action of that corporation or state or other non-human entity. Under this model, however, the actions of the non-human entity are necessarily always reducible to the action of a human being. Human beings take a central and special role as the only non-reducible actors. The impli-

52 Chopra & White 2011, pp. 2 and 154 respectively, although the point is made in different words on the respective pages.

cation here is that there is something about human beings that other types of entities such as corporations, states, animals, rivers or artificially intelligent entities lack and as such, it speaks of a particular metaphysical frame of reference. This approach is reflected in international law as well as the nominalist model of the corporation.

Under the other model, human beings do not hold the same special place in the metaphysical frame of reference, although human beings are still treated as prototypical. Some events caused by human beings are actions and some events caused by non-human beings are actions with no necessary link between the two. Non-human beings can be agents in the same non-reducible sense as human beings, and human beings can be actors in a reducible sense (as when there is an agent-principal relationship). The latter holds true, by the by, for the first model as well. This second approach is growing more prevalent. This is what French argues for with regard to corporations and what is reflected in the realist theory of corporations as well as the agential theory of states that is prevalent in international relations. Under this approach, genuine non-reducible action of non-human entities is taken to be possible; a non-human entity can act, provided it fulfils certain conditions that are taken to be necessary for 'genuine' or 'real' action.

To summarise, there are two variants of the metaphysical frame of reference currently reflected in the law. Non-human entities can act under both; in the one case they are 'only' capable of derived action reducible to human action, in the other case they are capable of non-derived action because they are in some way sufficiently similar to human beings. This takes for granted that human beings are non-reducible agents because they have certain characteristics, and that non-human entities can – at times – have the same characteristics. Under this view, it is those characteristics that make the entity an agent.[53]

In this context, two types of arguments regarding agency in law can be identified: on the one hand, there are normative arguments; on the other hand, there are metaphysical arguments. By the former, I mean arguments that recognise that regard the law as distinct from folk psychology or the metaphysical frame of reference and embed determinations of agency in a normative argument, for example in one regarding the problems regarding a kind of entity as a legal agent would solve and the desirability thereof. The normativistic approach of Neumann and Kelsen (cited above) reflects this. By the latter, I mean arguments that determine the question of agency by reference to the metaphysical frame of reference, depending on whether the entity has particular metaphysical characteristics, such as intentionality. French's approach (cited above) reflects this.

53 The normativistic approach, meanwhile, does not presuppose this, or it holds only that the conditions for 'genuine' or 'real' *legal* action are whatever conditions the law requires.

The example of animals being treated as agents as well as increased insights from philosophy and the cognitive sciences generally give us reason to question our metaphysical frame of reference:[54] it strikes me as important to carefully consider the metaphysical frame of reference that determinations of legal agency are based on, for the following reason: lawmakers (broadly understood) can in principle posit that legally speaking, an entity is an agent. If they do so by reference to the metaphysical frame of reference and it turns out that this frame of reference is flawed or incorrect in some way, this may remove the foundation for regarding that entity as an agent for the reasons that it has particular metaphysical characteristics. Let me give an example to clarify this point: I can choose to go to the train station daily at 3pm to pick up a friend who arrives at that time. If I learn my friend is not arriving at the train station at all today, I can still choose to go, but my reason for doing so is gone. Does it still make sense for me to go to the train station, then? Perhaps – but not for the reason that my friend is arriving, because she is not. I might go for different reasons, or I might have reason to go somewhere else (for example, if she tells me she will arrive at the airport instead). It seems to me that the principle is the same for law and what entities it treats as actors and for what reasons. Using international law by way of example:

> International law is currently positing that a State can only act when a human being has acted, and that action is attributable to the State for the reason that it takes it as a fact that only human beings can act ('properly', that is, in a non-derived manner).
>
> There are good reasons to believe that it is not a fact that only human beings can act ('properly', that is, in a non-derived manner).

What conclusion should be drawn from this?

The conclusion that international law should no longer posit that a State can only act when a human being has acted and that action is attributable to the State is too quick even when one ignores the logical invalidity of that derivation. What the above argument suggests, to me, is the following:

1. We should carefully investigate whether there are indeed good reasons to believe that it is not a fact that only human beings can act; that is, we should investigate and think about our metaphysical frame of reference. This is a task predominantly for the cognitive sciences broadly understood.
2. International law should not posit that a State can only act when a human being has acted, and that action is attributable to the State *for the reason* that only human beings

54 More on the role of philosophy and the cognitive sciences in Section 4.

can act, if it turns out that this reason appears misguided. This does not take away that it posits the same for different reasons. This requires normative arguments and is a task for legal scholars and practitioners, taking into account the insights from the cognitive sciences regarding (1).

4 Investigating the Metaphysical Frame of Reference

How can we investigate our metaphysical frame of reference and whether or not it is misguided that only human beings can act in a non-reducible manner and draw conclusions from this for legal agency? It seems to me that we need three things in order to do so successfully: firstly, we need to explicate the metaphysical frame of reference and investigate the metaphysical presuppositions that underlie the law, where such presuppositions are being made. Section II of this paper indicates that it is the case that such presuppositions are being made, but it does not constitute anything close to a full account of them, nor did I attempt to give such a full account in that section. Further investigation is required in this regard.

Secondly, we need insights from philosophy, neuroscience and the cognitive sciences more generally. Both philosophical argument and the cognitive sciences indicate that the division between subjects and objects, agents and things, persons and non-persons may not be 'out there' in an ontologically objective, mind-independent world. This entails that it is not the case that some entities can, as a matter of brute fact, act, while others cannot; that some entities are 'really real' actors and others are not. Intentionality, agency or personhood are not natural kinds, despite the fact that our minds may work in ways that suggest to us that they are via our intuitions.[55]

Some example can give us insight into what these fields can offer regarding the investigation of the metaphysical frame of reference. Farah and Heberlein suggest that

> [t]here is evidence from neuroscience and developmental psychology that these intuitions may come from innate brain mechanisms, separate from those used to learn about other concepts. Whereas the mental categories day and night, mountain and molehill are learned primarily by induction from experience with the world, we come into the world equipped with the tendency to categorize certain entities as persons and to treat them accordingly. Things with such

55 M.J. Farah & A.S. Heberlein, 'Personhood and Neuroscience: Naturalizing or Nihilating?', *The American Journal of Bioethics*, Vol. 7, No. 1, 2007, pp. 37-48; Greene & Cohen 2004.

> trigger features as humanlike faces and/or contingent movement and/or natural language will tend to elicit from us the 'intentional stance' (the assumption that the entity is acting according to its own beliefs and desires; Dennett 1978) and a certain moral regard.[56]

It seems that we attribute intentionality, agency or personhood (and many other things) to ourselves and to others. Dennett puts this in clear terms when summarising his take on the intentional stance and on intentional systems:

> An Intentional system is a system whose behaviour can be (at least sometimes) explained and predicted by relying on ascriptions to the system of *beliefs* and *desires* (and other Intentionally characterized features […]). There may *in every case* be other ways of predicting and explaining the behaviour of an Intentional system – for instance, mechanistic or physical ways – but the Intentional stance may be the handiest or most effective or in any case *a* successful stance to adopt, which suffices for the object to be an Intentional system.[57]

Taking this intentional stance toward an entity does not require or entail that this entity has any ''particular intrinsic features''[58] or that intentional interpretation of some entities are ''objectively true''.[59] It seems that our intuitions at times lead us to take the intentional stance also to stick figures or geometric shapes moving on a screen.[60] Moreover, a physical explanation for the behaviour of intentional systems may be possible in every case – including in the case of human beings.

Brozek and Janik indicate that what entities we take this intentional stance toward and regard as moral agents is culture-dependent:

> Moral agents must be recognised as such: to be a moral agent, an individual must be treated as a moral agent by other members of the community.[61]

---

56 M.J. Farah & Andrea S. Heberlein, 'Response to Open Peer Commentaries on ''Personhood and Neuroscience: Naturalizing or Nihilating?'': Getting Personal', *The American Journal of Bioethics*, Vol. 7, No. 1, W1-W4, at W2.

57 D.C. Dennett, 'Conditions of Personhood', *in* A. Oksenberg Rorty (Ed.), *The Identities of Persons*, University of California Press, Berkeley 1976, p. 179.

58 Id.

59 Id., p. 181.

60 F. Heider & M. Simmel, 'An Experimental Study of Apparent Behavior', *The American Journal of Psychology*, Vol. 57, No. 2, 1944, pp. 243-259.

61 B. Brożek & B. Janik, 'Can Artificial Intelligences Be Moral Agents?', *New Ideas in Psychology*, 2019, p. 3.

Moreover,

> to be recognised as a moral agent one has to fit within the concept of agency adopted in the given culture (and the folk psychology associated with it). Some cultures may consider animals or even inanimate objects (rivers, volcanoes, etc.) as capable of moral agency; others may exclude from the pool of moral agents some human beings (slaves, enemies, people belonging to other tribes).[62]

The folk psychology of a given culture and its concepts of agency, intentionality and personhood are all part of what I call the metaphysical frame of reference in this paper. Insights from the cognitive sciences broadly understood (including philosophy) can aid us in explicating and testing the metaphysical frame of reference that – at times – underlies the law and legal determinations of the agency of entities.

Thirdly, and in light of this and in light of the fact that law is *in principle* distinct from the metaphysical frame of reference and free to determine its own concepts, a normative theory is required. The normative theory is necessary in order to determine what the relationship between folk psychology and the life experiences of individuals (which are based on intuition and the metaphysical frame of reference), the law and insights from the cognitive sciences broadly construed should be. In other words, it is required to determine in how far law should take insights from the cognitive sciences into account or how closely it should reflect folk psychological notions of, for example, agency.

5 The Relevance of Investigating the Metaphysical Frame of Reference

If it turns out that the metaphysical frame of reference is misleading or incorrect,[63] and where the law refers only to the metaphysical frame of reference as reason for regarding an entity as agent, this takes away the foundation of that determination. Let me use international law as an example again:

> If it turns out that it is not an elementary fact that only human beings can act in a non-reducible manner, this takes away the foundation for international law holding that a State can act only when the acts of a human being are

62 Id., p. 4.

63 Whether 'misleading' or 'incorrect' is the more accurate term also needs to be investigated. This depends on whether it is a matter of truth in the sense of corresponding to objective (or social) reality, or whether it is something that is more or less useful, depending on its purpose.

> attributable and attributed to the State *because* it is an elementary fact that only human beings can act in a non-reducible manner. This does not mean that international law cannot hold that a State can act only when the acts of a human being are attributable and attributed to the State, but it takes away the reason why this is currently being held.

Moreover, the emergence of more and new kinds of non-human entities such as artificially intelligent entities gives a very practical reason to investigate the metaphysical frame of reference and the law's relationship to it: the question how to regulate these new entities will need to be settled and is already on the agenda of, for example, the European Parliament.[64] Questions such as whether these artificially intelligent entities should be regarded as legal agents, under what circumstances and for what reasons will need to be answered. I hope to have shown that we should not answer them on the basis of unreflective intuitions (probably) stemming from our metaphysical frame of reference.

6 Conclusion

In this paper, I have argued the following:

- Different non-human entities have been regarded as legal agents at different times. The conceptualisation of legal agents and legal acts underlying the choice as to which entities to regard as legal agents shows that this is – at times – done for metaphysical reasons: by reference to the metaphysical frame of reference. At other times, it is done on for normative reasons.
- There are reasons to investigate the metaphysical frame of reference.
- If it turns out that the determination of legal agency is made on the basis of references to the metaphysical frame of reference and if it turns out that the metaphysical frame of reference is incorrect or misleading in this regard, the reason for that particular determination of legal agency *on that particular basis* disappears. This does not take away that the same determination can be made for other reasons.

This – and the fact that the number of non-human entities that may need to be regulated is on the rise – leads me to conclude that the topic of non-human agency is a worthwhile subject for legal scholars and cognitive scientists to investigate. I have identified three (very broad) main topics that need to be investigated in order to successfully investigate non-human agency in the law:

64 *See* introduction.

1. Firstly, the metaphysical presuppositions that underlie the law need to be identified and explicated, where the law makes such metaphysical presuppositions. This is a task for legal scholars primarily and concerns a critical investigation of the reasons for which the law regards different non-human entities as agents.
2. Secondly, the metaphysical frame of reference needs to be investigated. This entails investigating whether the metaphysical presuppositions made by the law are correct, although these investigations need not be geared toward law specifically. I consider this a task for cognitive scientists broadly understood (and firmly counting philosophers amongst their numbers).
3. Lastly, a normative theory is necessary in order to determine for which reasons law should regard a non-human entity as a legal agent and in which situations the law should identify an event or situation as a legal act. This theory will also need to determine what the impact of the (moral) intuitions of individuals (based on folk psychology/the metaphysical frame of reference) or of insights from the cognitive sciences (that is, point 2) should be for determinations of legal agency. I consider this a task for legal scholars and legal decision-makers, taking into account the information under 2.

Insights from 1 and 2 are input that is required for the normative theory (3). Taken together, these three questions provide the foundation for a project on non-human agency that will – ideally – enable legal decision-makers to make well-founded choices regarding the agency of non-human entities such as corporations, States, churches, rivers or artificially intelligent entities.

References

Beirnes, P., 'The Law is an Ass: Reading E.P. Evans' The Medieval Prosecution and Capital Punishment of Animals', *Society and Animals*, Vol. 2, No. 1, 1994, pp. 27-46.

Brozek, B., 'The Troublesome Person', *in* V. Kurki & T. Pietrzykowski (Eds.), *Legal Personhood: animals, Artificial Intelligence and the Unborn*, Springer, New York 2017.

Brożek, B. & Janik, B., 'Can Artificial Intelligences Be Moral Agents?', *New Ideas in Psychology*, 2019, p. 3.

Chopra, S. & White, L., *A Legal Theory for Autonomous Artificial Agents*, University of Michigan Press, Ann Arbor 2011.

Cohen, E., 'Law, folklore and animal lore', *Past and Present*, No. 110, 1986, pp. 6-37.

Dennett, D.C., 'Conditions of Personhood', *in* A. Oksenberg Rorty (Ed.), *The Identities of Persons*, University of California Press, Berkeley 1976.

Ewald, W., 'Comparative Jurisprudence (I): What Was It like to Try a Rat?', *University of Pennsylvania Law Review*, Vol. 143, No. 6, 1995, pp. 1889-2149.

Farah, M.J. & Heberlein, A.S., 'Personhood and Neuroscience: Naturalizing or Nihilating?', *The American Journal of Bioethics*, Vol. 7, No. 1, 2007, pp. 37-48.

Farah, M.J. & Heberlein, A.S. 'Response to Open Peer Commentaries on ''Personhood and Neuroscience: Naturalizing or Nihilating?'': Getting Personal', *The American Journal of Bioethics*, Vol. 7, No. 1, 2007, W1-W4.

Fleming, S., 'Moral agents and legal persons: the ethics and the law of state responsibility', *International Theory*, Vol. 9, No. 3, 2017, pp. 466-489.

French, P.A., 'The Corporation as a Moral Person', *American Philosophical Quarterly*, Vol. 16, No. 3, 1979, pp. 207-215.

Greene, J. & Cohen, J., 'For the Law, Neuroscience Changes Nothing and Everything', Philosophical Transactions: Biological Sciences, 2004, pp. 1775-1785.

Heider, F. & Simmel, M., 'An Experimental Study of Apparent Behavior', *The American Journal of Psychology*, Vol. 57, No. 2, 1944, pp. 243-259.

Kelsen, H. & Wedberg, A., *General theory of law and state*, Harvard university press, Cambridge Massachusetts 1945.

Naffine, N., 'Who Are Law's Persons? From Cheshire Cats to Responsible Subjects', *The Modern Law Review*, Vol. 66, No. 3, 2003, pp. 346-367.

Newman, G., *The punishment response*, J.B. Lippincott, Philadelpia 1978

Roef, D., 'Corporate Criminal Liability', *in* J. Keiler & D. Roef (Eds.), *Comparative Concepts of Criminal Law*, 3rd edn, Intersentia, Antwerp 2019.

Schlosser, M., 'Agency', *in* E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2015 edn., https://plato.stanford.edu/archives/fall2015/entries/agency/.

Shapiro, S.P., 'Agency Theory', *Annual Review of Socioly*, Vol. 31, 2005, pp. 263-284.

Volk, K., LuDerssen, K. & Kempf, E., Unternehmensstrafrecht, De Gruyter, Berlin 2012.