



Universiteit
Leiden
The Netherlands

Reason Beyond Knowledge: A Contemporary Analysis of the Problem of Collective Action and Its Possible Solutions

Kerkhoffs, Marijn

Citation

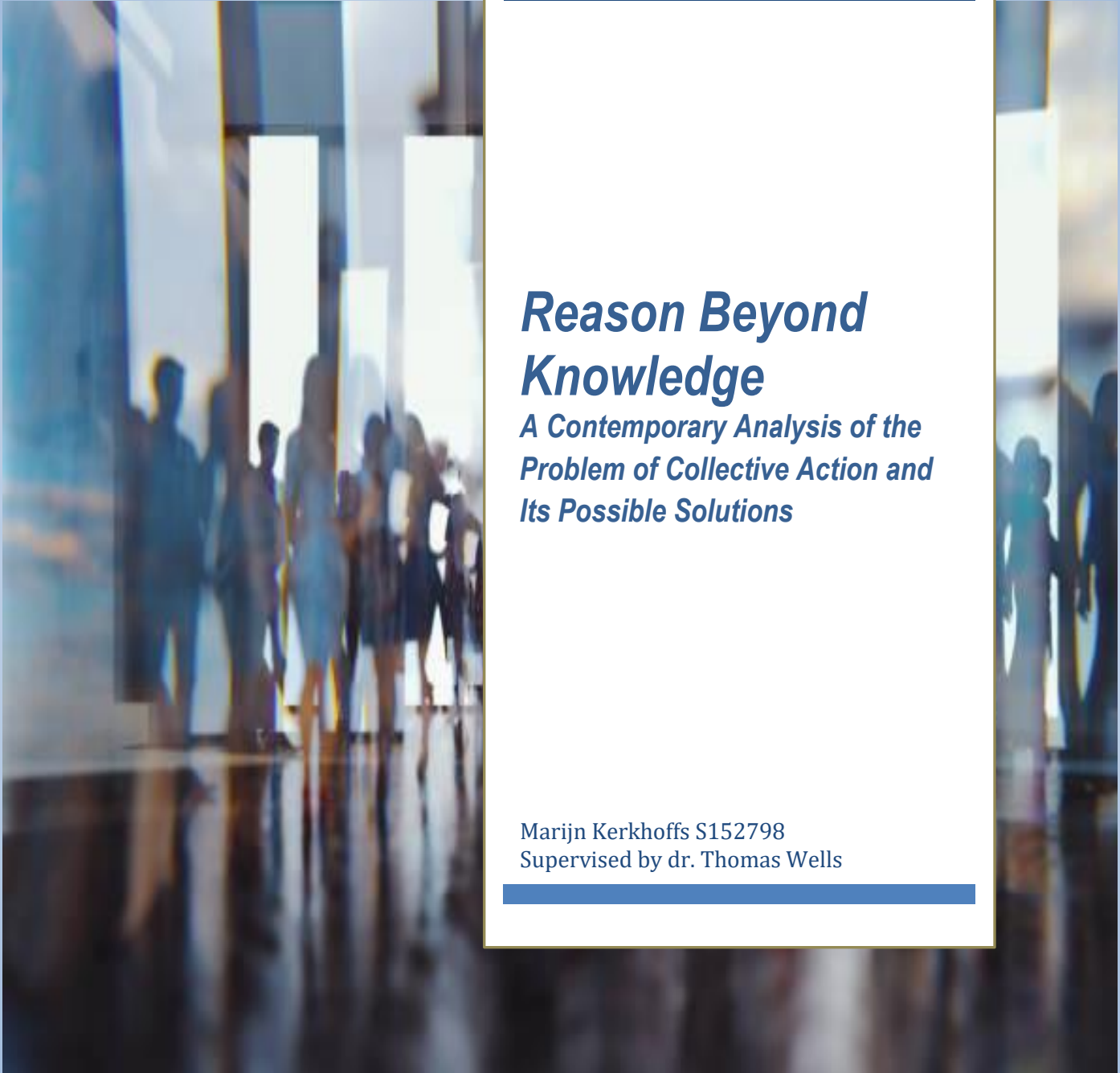
Kerkhoffs, M. (2025). *Reason Beyond Knowledge: A Contemporary Analysis of the Problem of Collective Action and Its Possible Solutions*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/4280936>

Note: To cite this publication please use the final published version (if applicable).



Master Thesis Leiden University –
Philosophical Perspectives on Politics
and the Economy

Reason Beyond Knowledge

***A Contemporary Analysis of the
Problem of Collective Action and
Its Possible Solutions***

Marijn Kerkhoffs S152798
Supervised by dr. Thomas Wells

Reason Beyond Knowledge: A Contemporary Analysis of the Problem of Collective Action and Its Possible Solutions

Table of contents

Acknowledgements	p.2
1. The Logic of Collective Action and Why It Is Problematic	
1.1 Contemplative Action	p.4
1.2 Contractarianism	p.4
1.3 Contractualism	p.5
1.4 Olson's Collective Action Problem	p.6
1.5 Conditions of the Two Problems of Collective Action	p.8
2. A Historical Perspective to Collective Action and the Prophecy of Its Demise	
2.1 Collective Action Through Religious Morality	p.11
2.2 Parallels of Morality and Collective Action	p.12
2.3 Collective Action Through Religious and Secular <i>As If</i> Morality	p.13
3. Collective Action As Moral Necessity: Deontology As A Possible Solution	
3.1 The Principle of <i>Ought Implies Can</i>	p.21
3.2 <i>Ought Implies Can</i> in Collective Action	p.21
3.3 Plausibility of <i>Ought Implies Can</i>	p.23
4. (Inter)personal Existence: Limits to Existentialism	
4.1 Existentialism: Existence Precedes Essence	p.26
4.2 Existentialist Morality	p.27
4.3 Problems of Existentialist Morality	p.28
5. <i>Homo Cooperatis</i> : the Existentialism and Essentialism of Human Being Revisited	
5.1 A Return to Essentialism?	p.31
5.2 The Ethics of Hybrid Existentialism	p.33
5.3 Problems of the Hybrid Existentialist Ethics	p.34
5.4 Implications and Considerations Concerning Hybrid Existentialism	p.35
5.5 Conclusion	p.38
Bibliography	p.40

I would like to thank my thesis supervisor, Dr. T. Wells, for his continuous support and the constructive sparring sessions during the long time this thesis was under construction,

my parents for their unwavering adamentness in times that I felt this thesis would never reach its conclusion,

my girlfriend Marlies for her incredible patience in providing moral support and academic motivation throughout the whole process,

and last but not least all the people that have helped me to become a better human than I was when I first set foot at this university

“... and I didn’t believe in magic, until I watched it disappear.”
– Gnarlis Barkley

1. The Logic of Collective Action and Why It Is Problematic

1.1 Contemplative Cooperation

As long as mankind has roamed this earth, it has done so in coexistence with others. Even the hermit once belonged to the social realm, and maybe still does, as he affirms it by his desire to be relieved from it. As with most species of animals, we see that mankind assembles itself in smaller and larger communities, each with their own rules and ends. Reasons to join communities are diverse, but some that come to mind are creation and affirmation of identity, discovery and creation of purpose, shared objectives, shared interests, or behavioural control. All these reasons appear to have an individual and a social aspect to them: individuals want to survive in the world and communities offer forms of efficiency and self-understanding that mere individuals cannot yield, thus it appears not only reasonable, but necessary for individuals to join communities or to maintain their relationships within the communities they find themselves in¹.

Similar to animal communities, a key aspect of cooperation appears to be physical survival. Imagine a single huntress trying to hunt a deer compared to multiple hunters doing the same thing. The former has to make all preparations by herself, has to teach herself how to make a bow and use it, and after the hunt, has to skin and cook the animal by herself. As she does all these things, the carcass begins to decay, leaving the huntress with little to enjoy from her hunt. With multiple hunters, tasks can be divided, animals can be tracked down faster, and hunts will most likely be far more efficient. Additionally, after the animal has been cooked, it can be divided amongst the hunters in time, making for an efficient division of labour (the hunting, skinning and cooking can be done by multiple people, without exhausting any of them) and benefit (the animal can be divided among a group of people, with no meat going to waste).

Some animals, however, seem to have some form of instinct that makes them cooperate because it is in their nature; human cooperation on the other hand appears to be based on individual contemplation and calculation, which leads the individual to trade off liberties for securities, mostly by agreeing to common boundaries, like norms, practices and laws. It is here that we find an odd situation, which forms the basis for contractarians like Hobbes or Locke: was there ever a time in which no cooperation existed and two individuals first agreed on something, and if there was, how did we get to the state of widespread cooperation we find ourselves in today?

1.2 Contractarianism

Both Hobbes and Locke propose an imaginative 'state of nature', which they then use to argue that humans have an inherent desire to leave it, hence the existence of communities and states. Hobbes argues that in a state of nature, any individual has the absolute freedom to do whatever lies in their mental and physical power, but they also have no security

¹ From birth, we grow up in certain communities that shape our identity and relationship with others humans and the world, thus some communities and cooperation are not actively chosen in the first place, although they can be repealed by active choice on most occasions.

beyond this individual power (Hobbes, 1929, 94-97). Locke argues that even in a state of nature, there would be natural laws that secure the right to life, property, and liberty to any individual (Locke, 2003, 102). Even though Locke's account of the state of nature appears far more preferable than Hobbes's in terms of personal security, both argue that mankind should – and will – find its way out of this predicament: Hobbes argues that individuals in the state of nature will contemplate trading in liberties for securities by yielding individual liberties to a state or ruler that can and will protect against the powers of its fellow men (cooperation) or a state or ruler will force him to yield liberties for securities (domination) (Hobbes, 1929, 131-132), whereas Locke holds that the basic liberties provided by natural law lead to the establishment of other conventions and codification of laws, which leads to the establishment of civil society (Locke, 2003, 141-143). Interestingly, both views imply what Helmuth Plessner has famously called the 'excentric positionality' of a person: unlike animals, human persons are not only aware of their position in the world, they can also look at themselves and their position from an imagined position in the world. By doing so, persons acquired a form of self-reflection through an in-the-world perspective (Plessner, 2019, 290-294). It is this self-reflection through a worldly lens that allows a person in the state of nature to contemplate that their chances of survival are higher when they cooperate with other persons, leading them to seek cooperation.

However, both theories appear unstable when it comes to their cooperative features. If we take Hobbes's account of a person in the state of nature, it appears unlikely that it will yield its liberties to form a small community, as Hobbes also holds (Hobbes, 1929, 131). How, then, does one person gain enough power to have other persons submit their liberties? And if we assume that the state of nature has never and will never exist, then what legitimacy does a ruler have in keeping liberties at bay, given that securities were never in peril? In Locke's case, we face another problem: Locke assumes natural laws that all persons are accountable to, yet without institutions, it remains unclear how people can be held accountable (something which is usually taken to be a basic premise for accountability to have any meaning in the first place). Locke appears to beg the question as to how society comes to be by postulating natural laws that, if they were to indeed exist in this way, would lead to the creation of civil society. Locke thus retorts to natural law being divine law (Locke, 1999, 336). But to use divine intervention as the cause of cooperation seems shaky, which we will further discuss in section two.

1.3 Contractualism

Modern supporters of contract theory have tried to evade contractarian problems by assuming a more deontological position towards cooperation. John Rawls argued in his famous *A Theory of Justice* that if persons had to discuss the workings of society without knowing their position in said society, it would be rational for them to come to two principles of justice that would be the basis of their society. Rawls supposes that these principles are as follows:

- 1.) Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.

- 2.) Social and economic inequalities are to be arranged so that they are both:
- a.) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
 - b.) attached to the offices and positions open to all under conditions of fair equality of opportunity. (Rawls, 1999, 266)

Rawls holds that with everyone being equal behind a 'veil of ignorance', they would make sure that whatever position they hold after the veil is lifted is still a position worth having. Given that this may be the lowest relative position for anyone, the lowest position must be as well-off as possible. In the same vein, all persons would agree to as many liberties as possible, given that some liberties will hinder other liberties.

As favourable as Rawls's position may look, this methodology, too, has faced critique. David Lewis Schaefer has argued that the original position of ignorance that Rawls posits makes it impossible for persons to unbiasedly deliberate, as particular knowledge is needed to contemplate positions in society, such as a knowledge of good and bad. Without this particular knowledge, it seems impossible to determine what the worst position in society may be (Schaefer, 1974, 94-95). Additionally, although Rawls does not assume natural laws like Locke, he still assumes that when the veil is lifted, men will still respect the principles of justice because they have a 'sense of justice' (Rawls, 1999, 125).

If we were to have such a 'sense of justice', would it be enough to make us uphold the contemplated principles of justice? Rawls argues that if it were not, then it may seem that the principles are not the right principles (Rawls, 1999, 125-126). But if this is the case, then we might well ask if any 'right principles' can be found, as it appears that the robustness of the principles depends on a real-world possibility of returning to an original position, or at least to a possibility of redistribution of positions not linked to historical ties and personal achievements. As such, Rawls's position appears too idealistic to account for the creation and upholding of civil society through cooperation.

1.4 Olson's Collective Action Problem

There appears to be something counterintuitive to the idea that persons will rationally opt for cooperation and security over absolute personal liberty. Both versions of contract theory propose some inherent sense of morality or justice towards other agents that should prevent us from acting upon certain personal liberties. However, even if we assume that we have such a moral sense, it does not seem decisive: I can still choose to act against Rawls's two principles of justice, for example, if I feel it benefits me more than acting per them. In society, especially when the expected consequences of non-cooperation do not weigh against the expected savings of non-cooperation, accumulated with the benefits of social cooperation upheld by others, this 'free-riding' appears to be the rational approach. However, if everyone were to 'free-ride', cooperation should be impossible to begin with. How then is it possible that we still have a functioning, cooperative society on the one hand, but individual incentives to cheat the system on the other?

This question is at the heart of what Mancur Olson has famously called *the problem of collective action* in his work *The Logic of Collective Action: Public Goods and the Theory of*

Groups. Olson argues that whether or not a group succeeds in providing a collective good depends mostly on three factors: the individual need of group members for the good, the chances of defecting being witnessed, and what clear consequences defecting has for those close to the defector (Olson, 1965, 43-45). He argues that the latter two factors are for a large part dependent on the group size: more people in a group generally means individual actions are less scrutinised and individual free-riding has less influence on the success of collective action, which in turn means less responsibility for a collective good not being provided to those in need of it (Olson, 1965, 45). The first factor means that sometimes, collective goods may be provided without collective action when one individual values a collective good so much that they want to provide it for themselves (and with that for others) even if they have to pay the full cost (Olson, 1965, 44, 46).

We can summarise Olson's argument for the possibility of collective action using the following reasoning:

Collective goods can be provided without coercion or incentive iff:

1. One or some individuals value the collective good it yields so highly that they wish to provide it for themselves, effectively providing it for others as well, or
2. Individual contributions make a noticeable difference (to at least some agents) to the success or failure of the collective action needed for the collective good, or
3. Failure to contribute noticeably leads to failure of the collective action needed for the collective good and consequently to noticeable harm to those who have to go without the otherwise provided collective good.

As the first factor focuses on the provision of a collective good by means other than collective action, it is outside the scope of our investigation. However, if we scrutinise the other two factors, a key feature stands out: both factors seem to have a hidden premise. If our failure to cooperate is being noticed by other agents or the effects of collective failure on others are noticeable to us, what would that entail? If there are no consequences like punishment or reward, why would we still be inclined to cooperate in these circumstances?

It appears that the proposed factors are necessary, but not sufficient to explain the workings of collective action. They are necessary because given that one would be rational, even if one would be selfless, it would still be irrational to contribute to collective action given that both the impact of the contribution nor the impact of provision of the collective good to any agents would be noticeable (Olson, 1965, 64-65). Olson seems to correctly hold that social pressure and social incentives affect smaller groups more than larger groups (Olson, 1965, 62), but this cannot explain why smaller groups may succeed in acting collectively without *any* social or economic incentives.

One may ask why this matters. After all, there seems to be plenty of social and economic incentives around for us to motivate us to act collectively. Why would we need the possibility of collective action without these incentives? To understand why the lack of such collective action yields a problem, we need to take a closer look at the incentives that may motivate us to act collectively and where they come from. Let us, for example, take a law that forbids us

to litter. If we are caught littering, we will have to pay a fine. If the risk of being caught and fined is greater than the benefit of disposing of one's trash without having to look for a bin, we should be motivated to abide by the law in question. However, we will only get caught if there is law enforcement. If we then ask how law enforcement comes to be, we find ourselves in a similar situation: it is only through some form of collective action that a police force is established. We can say the same for the law in question: without some collective deliberation and agreement, it could never come to be. It seems the problem of collective action yields two specific problems: firstly, we may ask how we got to any collective action to begin with if every collective action needs other collective actions to work. We can call this the regression problem of collective action. Then, we may also ask how existing collection persists when other forms of collective action cease to persist. We can call this the erosion problem of collective action.

1.5 Conditions of the Two Problems of Collective Action

Having established two different but related collective action problems, we can now turn to see how these problems may come to be and if they are as adamant as they appear. Olson holds that humans are rational and self-interested beings, which means that they will only act collectively when their interests are better served through collective action and there is a reasonable chance that the collective action will succeed. Additionally, if there is a reasonable chance the collective action will succeed without their support, they will also refrain from contributing since that position is then likely to yield extra benefits without costs. But, Olson also argues that, even if we were not self-interested but altruistic instead, it would still not be rational to contribute to collective action given that there is either no reasonable chance of success or if a contribution will not lead to any noticeable difference.

It thus seems that it is the rational component of human existence that is crucial in explaining why we face the two problems of collective action. This also makes sense if we recall Plessner's notion of excentric positionality: animals have no excentricity, which means that they cannot reflect upon their position in the world as if they were in-the-world instead of in-themselves. This means that all collective action for animals is either immediate or non-existent: an animal will act on an innate instinct to cooperate,² or it will not cooperate at all³. Human excentric positionality proves both a challenge and a gift, because we can reflect on our position in the world, and this reflection becomes a medium towards cooperation or non-cooperation. We may actively decide whether we find a cooperation to be worth our while. Given that we can reason that cooperation will only succeed if others cooperate as well, this means that we need external motivation in the form of some certainty that others will do so. The gift of our predicament would be that we can accomplish more complex cooperation with larger groups and greater benefits, given that we can establish and maintain the necessary conditions for successful cooperation.

² Or instinctual motivations that are necessary for survival, like acquisition of food, security and reproduction.

³ It must be noted that in recent years and in specific animal groups, like apes or dolphins, new forms of cooperation have been recorded. It would thus seem that either new forms of cooperation can be developed if they are useful for the instinctual motivations, or (some) animals may also possess a form excentric positionality.

Since Olson has shown that it makes little difference if we were to be inherently altruistic or self-interested, we are left to ponder whether inherent⁴ human rationality implies non-cooperation in all circumstances, and if inherent human rationality is plausible. James Q. Wilson has argued in his *Political Organizations* that Olson falsely claims rationality and self-interest are homogenous terms: one can only have one ultimate rationale or interest, competing rationales or interests will eventually be dismissed for one's true rationale or interest (Wilson, 1995, 23). Wilson argues that we can act rationally and irrationally at the same time for different interests, for example when we donate to charity: while it may be rational for me to do so because it makes me feel morally good, it may be irrational to do so in a socio-economic sense, since my singular donation does not make a perceptible difference to anyone (Wilson, 1995, 23). Wilson thus argues that Olson creates a problem by arguing that the rationale of self-interest is socio-economic by definition. If we consider other rationales that may be self-interested, like wanting to feel righteous, we may contribute to collective action on irrational socio-economic grounds and still act rationally self-interested (Wilson, 1995, 24).

However, we may question this reasoning by asking whether non-socio-economic rationales exist as ultimate ends or whether they are means to socio-economic ends. In the donation example, we may well argue that the reason we feel good while donating is because it gives us a certain sense of (moral) worth. As such, it seems that there is still a hidden socio-economic transaction going on, which would make my self-interest ultimately socio-economic. We could only examine whether truly non-social value exists if no social structures or values existed, and one would still feel inclined to donate. We would then fall into a problem of progression: if no social structures or values exist, how could collective organisations like a charity exist?

Another critique of Olson comes from David Gauthier, who has argued that even if the notion of rational self-interest is inherent, it does not imply that we cannot come to collective principles. He argues that we can imagine two human states of being after we have identified inherent rational self-interest: we are either unconditionally non-cooperative (which Gauthier classifies as "egoistic") or conditionally cooperative (Gauthier, 1987, 2). A third state, unconditional cooperation, can be dismissed because it seems to contradict the notion of rational self-interest, as necessary cooperation would mean that we do not get to choose whether or not we act out of self-interest. This means that the two other states remain possible. Now, the rational self-interested individual in the egoistic state would have fewer possible outcomes when it comes to maximising its self-interest, given that they can imagine other individuals being in the conditional state. As such, it would be in its self-interest to be in the conditional state, given that they believe they can properly make out who else is in this state while facing a collective action problem. This leads Gauthier to believe that rational self-interested individuals will be conditional cooperators and thus, collective action should be able to be established even within a group of rational self-interested individuals (Gauthier, 1987, 9, 15, 61).

Nevertheless, collective action being possible does not mean that it will also come to be. Gauthier's reasoning seems to rely for a great part on the premise that individuals can make out whether or not others will cooperate. This means that individuals must assume that the

⁴ What we mean here by "inherent" is "inescapable, necessary to be considered human"

fact that they can think of collective outcomes must also apply to others, and that others also think of the individual in question. In any other case, the term conditional cooperation appears to be an empty phrase, as individuals are conditioned not to cooperate, leaving them in a state of egoism in practice. As we have seen in Olson's reasoning, the effects of singular contributions to collective action need always be noticeable for individuals to contribute. This appears true even for conditional cooperators, as they can rationalise that if their contribution will not be noticed, others will realise the same and will not contribute, as their non-contribution will surely go unnoticed. Additionally, the notion of individuals being aware of cooperative outcomes seems to assume that they have witnessed such outcomes. This can only be true if cooperation already exists before the conditional cooperator is faced with a collective action problem. Gauthier's argument cannot explain how there can be cooperation before individuals change their stance from conditional non-cooperation to conditional cooperation. These considerations mean that Gauthier's argument appears little convincing at the moment, but we will return to it in the last chapter.

Before we move on to the analysis of how collective action may still have come into being despite the regression problem and why it might collapse over time, it is important to note a few things we have found so far, as well as to give a concise overview as to what we need to consider in the next chapters. Our newfound collective action problem is not the same as the version described by Olson. For Olson, the collective action problem is a problem that springs from our self-interested rationality, which can then be solved by external motivation. The problem we have identified challenges the idea that external motivation can solve Olson's problem: infinite regression means external motivation cannot be a necessary component of successful collective action, given that collective action can succeed at all, so either self-interested rationality in itself can give reason to cooperate or there needs to be something else that can bend self-interested rationality in a way that collective action can still succeed. This is not a mere logical or philosophical problem; rather, it touches upon the very foundation of our social existence: many economic and social benefits in society can only be achieved through stable cooperation. Stable cooperation requires us to be able to trust others to keep their promises and vice versa. If any person can identify external motivation as non-binding, anyone can choose to free-ride, which leads to steady erosion of all cooperative schemes.

In this thesis, we will compare different approaches to solving the collective action problem at hand in a dialectical manner, showcasing their vulnerabilities and strengths. The two main goals are to better understand the workings of the problem we have encountered and collective action as a whole, and to ultimately attempt to formulate a possible solution to it in the last chapter of this thesis.

In the next chapter, we will consider the possibility of a workaround for our concerns regarding external motivation, after which we will move on to the possibility of internal motivation as a solution to our problem of collective action.

2. A Historical Perspective to Collective Action and the Prophecy of Its Demise

2.1 Collective Action Through Religious Morality

In the last chapter, we have found that the solution to the collective action problem that Olson proposes creates its own collective action problem: if collective action always relies on external motivations like punishment or reward, these motivations being the result of earlier collective action appear unstable as we are faced with a problem of infinite regression. A possible solution to this regression would be to define a starting point that does not rely on collective action. If we briefly recall Hobbes's and Locke's contract theories, we may observe that both rely on some divine providence that guides us towards collective action. In this chapter, we will consider the possibility of divine providence as a primordial motivation for collective action.

First, we may consider a situation in which there is a common belief in a divine force, such as a god, that watches over us and specifically over our actions and promises towards other individuals. Additionally, this divine force can inflict some form of damage on an individual who does wrong to another individual. If these premises were commonly held to be true, then we find that such divine motivation could serve as an external motivation which, at first sight, needs no initial cooperation to be enforced. As such, the collective action problem would be solved by an unmotivated motivator⁵. Historically speaking, such a situation has been prevalent nearly everywhere. Even today, religious people still make up a great majority of the world's population. Two questions thus spring to mind: firstly, is a belief in a divine force a necessary condition for successful collective action, and secondly, is a belief in a divine force a satisfactory condition for successful collective action?

Let us recall the potential problems we found for a fundamental initial cooperation, since a fundamental superhuman external motivation seems to create a similar situation. We argued that if this were to be the case, all other forms of cooperation could effectively be derived from it. We might then say that all cooperation can be explained through this superhuman external motivation. This, however, makes for two problems that resemble each other closely: first, the ways and threats of the divine force must be interpreted by humans to establish coherent rules about how we should act to avoid punishment and how we expect others to act as well. Only then do we have a common ground of trust in others⁶ to do their part when it comes to collective action. The second problem we may encounter is a situation in which rivaling schemes of collective action may exist. It appears logically impossible that both can be in line with the ways of the divine force, so one of them has to be based on a wrong interpretation. The superhuman external motivation thus appears to assert little pressure from itself, since only its fallible interpretations can motivate us to cooperate. Only if we believe an interpretation to be truly infallible (dogmatic) does the desired external motivation hold.

⁵ Or at least by a motivator that is unmotivated by external motivations, but just 'is' in the way it motivates. We may even speak of internal motivation by proxy: humans are not internally motivated to do good, but God is and it exerts enough pressure to form an external motivation for humans.

⁶ Or trust in the divine force that it asserts enough pressure onto others to do their part.

As such, the position of divine force as a satisfactory condition for successful collective action seems to be in contention. However, it would be too simple to suggest that only a belief in a divine force or one specific interpretation of that force leads individuals to cooperate. If we look closely at religious pressure on, for example, morality, we find that it is not only a fear of divine punishment that leads people to act morally, but also the belief that other humans may also punish us based on divine law or contingent laws that are derived from it. But it seems that all these forms of external motivation still serve as mechanisms to make the self-interested rationality of individuals comply with the general goals of communities or societies: we do not act cooperatively out of altruism or trust in cooperation of others, but out of fear for harsh punishment and the assumption that others feel the same way. This means that for now, Olson's theory on collective action appears to hold, and the problem of infinite regression we have identified as the true problem of collective action does as well.

2.2 Parallels of Morality and Collective Action

Before we move on to discuss the problems that may arise when a belief in a divine force or interpretations of it falter, it is important to briefly return to an important notion. We used the example of morality in the last paragraph as something religion can put pressure on to make it work. But what exactly is morality, and why does it relate so closely to collective action? Since this question is very delicate and complicated, while also reaching far beyond the scope of this project, we will only briefly cover the role of morality concerning collective action. What is usually understood as morality is the will to do what is just. This definition itself, however, needs its terms to be well defined. When can we speak of a "will" to do what is just? And what can be identified as "just"?

Here we find a close relationship with collective action. Remember that we have made a distinction between instinctive collective action, like pack tactics in wolves, and voluntary collective action, which is the subject of our investigation. For morality as well as collective action, it seems that we only find it to be 'true' if it is not forced upon an agent. In case of collective action, we may say that external motivation to act cooperatively may render the cooperation untrue or unstable: as soon as the agent finds a better opportunity for itself, it may abandon its promise to cooperate. In the case of morality, we may say something similar, but maybe even more pressing: if external motivation is needed to act morally, can we even speak of morality in the first place, or should we speak of pseudo-moral opportunism?

The question of 'justice' appears somewhat trickier, since we can argue it to be either based on convention or some higher duty or feeling that would inherently be part of our being as humans. However, here too we find a convergence between morality and cooperation, since our project on cooperation tries to answer a similar question, namely, whether cooperation comes from internal motivation (duty, moral feelings) or external motivation (convention, fear of punishment). Given these similarities, we will treat morality as a part of the collective action problem from this point onwards.

2.3 Collective Action Through Religious and Secular As If Morality

Now that we have argued that 'true' morality does not seem to be able to have its roots in external motivation, we can also state that religious pressure does not make for 'true' morality, but for an *as if* morality or pseudo-moral opportunism: we act morally because it is more individually beneficial to act this way than to free-ride⁷. Fred Hirsch makes a similar observation in his *Social Limits to Growth*: the act of altruism through morality becomes *as if* altruism when *as if* morality is in place: we do not act cooperatively because we want to help others or because we trust others to do the right thing by their own accord, but because we feel it is necessary for self-preservation and we believe others to feel the same (Hirsch, 1976, 139, 146-147). We can see that Olson and Hirsch share their feelings about individual human motivation, and Hirsch feels that the external motivation that Olson's theory needs to solve the collective action problem has historically been generated by religious pressure (Hirsch, 1976, 138-139). However, Hirsch argues that religion, especially in the Western world, has been severely subjected to scepticism. As we have argued before, when it comes to divine force, dogmas appear to be a necessary component for further collective action. If individuals find these dogmas (or even the divine being itself) to be implausible, they are more likely to break with social norms and values that follow from religious institutions to gain more personal benefits (Hirsch, 1976, 117-118, 137-139). This movement would speed itself up over time with the reinforcement of religious institutions, withering as people cannot trust other people to fear retribution anymore, bringing back the original prisoner's dilemma.

Hirsch thus argues that over time, *as if* morality will erode. He also adds that this erosion may speed up in a capitalist society, since capitalist values such as personal wealth acquisition and competitive markets are effectively limited by (pseudo-)moral values like altruism, stewardship, or beneficence. To maximise the capitalist values, it is thus necessary to abandon these religious or (pseudo-)moral values (Hirsch, 1976, 117-118). We may add that if this were to be true, the same would be true for collective action: in a world in which personal gain is of utmost importance, there is overwhelming reason to believe that other agents are competitors before anything else, which makes it so that collective action cannot be achieved without external motivation, leading us yet again back to the problem we derive from Olson's solution.

However, if the loss of *as if* morality in individuals to secure personal benefits at the expense of cooperative action and its benefits were to be imminent, it would reasonably lead to less cooperation, fewer social norms, and eventually, fewer overall benefits. In other words, if, as Hirsch reasons, capitalism would be the default state of society given the egoistical nature of humans, this society's members would rapidly lose their ability to cooperate, leading to a decrease in overall welfare. Additionally, the loss of religious morality should lead to the rule of law being omitted (since it is based on cooperation), creating a state in which everything would be permitted for everyone. But, if we look at the economic and civil course of the secularised west over the last few centuries, the opposite appears to be true: as Steven Pinker argues in his *The Better Angels of Our Nature*, the rule of law appears to be stronger in current secular states when compared to their historical religious counterparts, resulting in less violence and casualties.

⁷ Or, in any case, the individual risks of free-riding are too high.

He attributes this trend to what he calls the ‘Civilizing Process’ and what is generally called the ‘Humanitarian Revolution’. (Pinker, 2011, 12). The former signifies the creation of central authority in the form of governments, the latter signifies a turn from religious morality to humanist morality grounded in reason instead of belief. Of course, the establishment of states and the growth of their power over the last five hundred years coincided with religious belief for the larger part. However, the legitimacy of states appears to shift at the end of the 18th century from a ‘god-given right’ to ‘a need to protect the commonwealth and its members’ (Pinker, 2011, 115). Pinker extensively describes the drawbacks of religious influence on the workings of the state, as he feels religious laws are dogmas arbitrarily derived from an interpretation of religious morality, legitimising religious violence against people believed to be heretics to protect the religious community (Pinker, 2011, 114, 119-123).

As we have already noticed, the workings of the state and those of religious pressure are fairly similar, except for the appeal to eternal life or suffering which some religions hold as part of their belief in a divine force: since religion only seems to generate *as if* morality, believers seem to act morally only as long as it suits their survivor interests. This may, of course, mean that some will act morally independent of how they expect others to act because they are fearful enough of eternal damnation through the divine force, but the aspect of others’ belief still plays a vital role, which we can illustrate with a scheme of a prisoner’s dilemma:

	Individual believes in religious morality (or laws derived from it).	Individual does not believe in religious morality (or laws derived from it).
Society believes in religious morality (or laws derived from it).	Everyone is limited by religious morality, equal freedom among members.	Pressure on individual to adhere to religious laws, fragile equal freedom among members.
Society does not believe in religious morality (or laws derived from it).	Individual limited by religious morality, more freedom for other members. Individual risks being exploited.	No one is limited by religious morality, infinite freedom among members, no cooperation possible(?).

If an individual acts under religious morality, but others do not, they (the individual) will end up in a position of relative disadvantage. For example, if I believe that I am morally obligated to keep my promises, but the person I am bargaining with does not, I will most likely be fooled into contributing while the other person does not. Of course, as we have also concluded from Wilson’s objection to Olson, this relative disadvantage does not have to go against my self-interest, since I may believe that my honesty will be rewarded in another way. However, it is very likely that if those holding on to religious morality continuously end up in relatively disadvantageous positions, they will eventually abandon their belief in religious morality, given that they are unsuccessful in persuading others to act as they do.

However, we would expect social norms and laws in a secular state to erode similarly, since the state and its laws are also schemes of collective action that stand and fall with compliance and reinforcement through behaviour of others and ourselves, which in turn would need external motivation generated by collective action. Pinker notes that states only

persist when their impositions and laws are seen as legitimate by the populace (Pinker, 2011, 72). He argues that what makes state impositions different from communal norms is their consistency and impartiality (Pinker, 2011, 73). The question we are then bound to ask, given that states would be less susceptible to erosion of morality or collectivity, is whether states are indeed as consistent and impartial as Pinker believes them to be.

As we have previously found, states, their governments, and their impositions are still forms of collective action. Collective action takes place, whether it be actively or passively⁸, through individuals. If Olson and Hirsch are right about the egoistical default state of human nature and consequently about capitalism being the default state of society were it not for religion, then individuals who engage in the political process will most likely enact policies that suit their personal goals. This means that with a limited number of political agents, a government could never be impartial, since the few would use their political power to further their own goals at the expense of others. Likely, such aristocratic, meritocratic, or oligarchic states would also be inconsistent, as it would be unreasonable for those in power to apply the same rules to themselves as they do to others. Could a democratic state successfully incorporate Pinker's conditions?

Hirsch argues that this is not the case, since the upkeep of social institutions like a government would require individuals to contribute part of their benefits (Hirsch, 1976, 148). Until now, we have found that only external motivation can ensure that rationally self-interested individuals contribute to such a scheme, with one important motivation being the belief that others will do so too. Any individual reasonably finds that it can dodge its duty to contribute as long as it is not caught. It can then reason that others may reasonably do the same. Even though such a state may thus be impartial because everyone has a say, its only consistency would be that all agents should reasonably disengage. As such, democratic governments thus should fall victim to the same circular regression of external motivation we have seen before: governments are a form of institutionalised collective action, agents have an incentive to secure personal benefits instead of upholding social institutions (collective action), social institutions lose their compulsive power as the amount of contributors decreases, which in turn leads to people having even less incentive to uphold the institution, etc. (Hirsch, 1976, 118, 132).⁹

How could it be, then, that society has still thrived for hundreds of years after the Humanitarian Revolution? Hirsch believes that it is the remains of religious morality that have kept individual egoism at bay, but that the pressure from human nature and the capitalist value system have finally deprived economics of moral values (Hirsch, 1976, 119-120).¹⁰ Why is this specific division of *as if* morality and economics relevant? Can *as if* morality not persist if it is not linked to economics? We may argue that it is indeed so that this division can speed up the demise of *as if* morality as a whole. Let us reminisce about how *as if* morality is the result of rational self-interest, I act morally, because it will benefit me

⁸ What we mean here by "passive collective action" is similar to the notion of tacit consent. We will further elaborate on this in the last chapter of this essay.

⁹ This line of argument has often been used against the libertarian idea that a society can govern itself based on individual incentives. It appears libertarianism too faces the problem of collective action when it tries to maintain society without collapsing into a Hobbesian state of nature.

¹⁰ An idea which resounded in economic thought at the time, as economists like Milton Friedman emerged as strong proponents for 'value-free' economics.

more than not acting morally. This argument is a classic example of an economic cost-benefit analysis: I gain (economic good X) if I perform (moral action A), but I will not receive (economic good X) if I perform (immoral action B). By taking morality out of the economic equation, there are no longer economic gains to be had through specific moral actions¹¹. All actions in the economic realm appear to now be amoral. In the initial cost-benefit analysis, there may have been a hidden value attributed to the morality of the action as well: if I perform (immoral action B), I may feel terrible about it. The shift to overall amorality in economics makes it so that even though (immoral action B) may appear morally wrong, it can never be so in an economic sense. As such, it appears to loosen (and eventually lift) the moral limits to economic acquisitions. Morality devoid of economic benefit, given that Olson and Hirsch are right, would be little more than something of personal value that holds back the individual in its pursuit of individual benefits. Once again, this can be self-interested still, but it is irrational in a socio-economic sense to hold onto and eventual abandonment seems immanent.

However, the strange implication of this dichotomy of the economic and social realms is that it seems to imply that after the social realm ceases to exist (because as *if* morality loses its meaning), the economic realm will become the social realm, as people will only act socially if it makes for personal economic benefits. But, as we have already concluded, there will be no reason to act socially at all because of the collective action problem. Effectively, this would mean that capitalism, if it were to become the social ideology, would mean the end of social activity, which it, in turn, needs to exist for capitalism to exist (after all, markets, competition, and even monetary value is still based on shared definitions and conventions). Thus, with the dichotomy of the economic and social realms, it seems capitalism would undo its social foundation. Hirsch notices similar events, but he ascribes the push for the dichotomy to liberalism as well:

“But in doing so, liberal philosophy had an unintended side effect. It undermined its own mechanistic instrument for attainment of individual goals. For the efficient working of the market itself rests on certain aspects of social morality that are affected by the means and motives prevalent in the economic system. As capitalism has become more mature and more managed, the stresses resulting from the social dichotomy have grown.” (Hirsch 1976, 124)

Now, liberalism is a very wide-ranging ideology with different interpretations. We will not go into depth on which version is meant or if the failure of collective action ought to be attributed to liberal philosophy. We can, however, take some general features of liberal philosophy to support Hirsch's claim. Two key features of liberal philosophy seem to be, on the one hand, a focus on universal rights and on the other, on individual liberties. These two features may contradict each other: if there is a universal right to life, I cannot, at the same time, have a liberty to kill someone. Given that a body of universal rights is established, it must be so that individual liberties are generally limited. Another limit to individual liberties would be the level of welfare one has. For example, given that I don't have a universal right to having a jacuzzi, I may only enjoy the liberty of having one if I have enough money to purchase one. It is thus important to grow individual welfare to increase the actualisation of individual liberties.

¹¹ At least in the long run, as to deny economic benefits on the basis of morality would become irrational.

Capitalism offers a way to do this by creating a market system in which any individual can obtain what they desire, ideally at a fair price based on supply and demand. With the openness of such a system, individual desires become clearer and can be met more easily as more demand should create more supply, lowering prices and thus growing individual liberty to fulfil specific desires.

What remains, however, is a tension between universal rights and individual liberties: if a universal right to housing exists, it would be an immoral situation if there are homeless people, so the housing market must be restricted to provide this right, resulting in less individual liberty to invest in the housing market or profit from it. The crux of the situation lies in the prevalence of either rights or liberties, given that both are part of the same socio-economic realm. By creating a dichotomy between economics, in which liberties are prevalent, and sociology, in which rights are prevalent, it appears liberalism, if it were responsible for this movement, has tried to have its cake and eat it as well: rights would still be in place, whereas liberties could expand limitlessly¹². However, as Hirsch argues, since there is only an *as if* morality (which is still based on economic benefits) instead of a real morality (based on natural duty or feelings that command one to do what is right), the rights deprived from their economic meaning lose their social meaning as well (Hirsch, 1976, 120-121).

What logically follows is a situation with no rights and endless liberties, theoretically speaking. In reality, however, some liberties can only exist if certain rights are in place¹³. If I have no right to property, it is impossible to enjoy the liberty of having a jacuzzi, at least not with the security of knowing no one will take it from me when I am not looking. If Hirsch is right, it is precisely here that we find the contradiction of the dichotomy between the social and economic planes: capitalism can thrive only if people believe they have a property right (otherwise they cannot strive for personal benefits), but by excluding any right from the economic realm, one must find personal benefit to the right to property or come to a social convention apart from economic benefits to all contractors. As we have seen, the latter appears impossible given that only *as if* morality exists. Without collective action to protect and enforce such a right, one will lose interest in upholding it for oneself, since others will breach this now mere individual right at any moment it provides individual benefits to them. So, the dichotomy would mean that eventually capitalism - and liberalism as well - would become impossible to sustain.

Hirsch sketches a gloomy situation in which the *as if* morality is slowly but steadily abandoned as 'moral' institutions like the law or social norms become subject to questions of economic legitimacy, which they do not have by definition since they are outside of the economic domain. With no economic reason to sustain these institutions, especially with moral control being relatively absent in a greater population of purely individualistically driven citizens, they will eventually wither completely (Hirsch, 1976, 131-132). With morality dissolving under the pressure of economically driven values, the economic system itself faces the pressure of a purely economic calculus:

¹² Socially speaking, of course. There would still be material and spatial limits.

¹³ This is generally the idea of positive freedom as introduced by Isaiah Berlin in his work *Two Concepts of Liberty*.

“Why should I adopt moral standards helpful to the system if the outcome of the system for me cannot be validated on moral criteria? True, the system is said to work out for people as a whole, compared with the alternatives. But I am not people as a whole, I am me; and unless the system can be shown to give me a fair deal in the only currency it deals in —material advantage— it can't ask me moral favors.” (Hirsch, 1976, 134)

The moral favours Hirsch refers to are those needed for a capitalist¹⁴ society to function, such as being true to one's promises and trusting other people to be true to their promises as well (Hirsch, 1976, 120-121). As we have already found, without these premises, collective action in any form appears impossible due to the collective action problem.

However, where Olson believes the collective action problem seeps from the economic domain into the social domain, Hirsch believes that this may be a step too far. He notes that the two premises Olson makes, namely that individual objectives are always directed towards private goals and that individual behaviour is always in line with these objectives, do not work as rigidly in practice as they would in theory: if this were to be inherently true, no social rules or conventions could receive any support, because they would take an initial sacrifice from any contractor, which would be irrational according to Olson's principles (Hirsch 1976, 136). Yet, we still see many forms of social convention and support for social rules in society, which would have to mean that either the first premise is wrong and people ultimately value other things than private benefits or the second premise does not follow the first one (because either human behaviour is not (fully) rational, or because there is an intellectual disconnect between human desires and the human behaviour necessary to fulfil them).

Hirsch believes the first premise to be false, since he does not believe that individual objectives are necessarily directed towards private goals. He mentions that already existing group values and group processes influence individual objectives towards more public goals (Hirsch, 1976, 136). This idea of inherent group thinking in individual humans is an important part of the communitarian view on collective action, which we will return to later. A question we may well ask at this point is whether these group values and group processes are not part of schemes of collective action as well and, as such, should be vulnerable to gradual loss of enforceability over time in the same manner as laws and norms would, given the loss of *as if* morality. Hirsch does not give an argument as to why these group tendencies are immune to the collective action problem we have posed, and as such, it seems that Olson's proposed premises still seem to hold for economic as well as social situations.

Hirsch then proceeds to argue that to maintain cooperation in society, we need either compulsion or an internalised social ethic (Hirsch, 1976, 144). We have already argued that compulsion can only be enforced through cooperation itself, which makes it vulnerable to the collective action problem. Hirsch believes that by internalising a social ethic, for example through *as if* morality, we may bypass the collective action problem by creating social norms that reinforce themselves. It is important to note that there is a major difference between internal and internalised, as the former means that something is a natural given, whereas the

¹⁴ Or quite possibly any society.

latter is a convention that is treated as being naturally given¹⁵. Hirsch holds that internalised social ethics may 'override' the rational self-interest that Olson thinks to be natural (Hirsch, 1976, 147). He stresses the need for internalisation instead of mere acceptance of external social motivation because the default state of human rational self-interest would mean that the latter would be unstable:

"These expenditures have a cost in terms of other opportunities foregone, so the operative question is whether the benefits —summed in some way over all the individuals to whom they accrue— exceed the costs. Since each individual can accurately measure only the benefit to himself or herself, the relevant entity for purposes of this information is the individual's own private gain derived from the facility —the amount that he or she would be willing to pay to have the public good available for personal use. The snag is that to obtain this information and to hold individuals to paying up the equivalent of their benefit or some proportion of it to finance the collective project will be impossible while individuals seek to maximise their individual gains, since they will then have the incentive to dissemble. If you pay what you say it is worth to you, you have an individualistic incentive to say it is worth little. Even the less ambitious task of collecting compulsory taxes for the public project on some objective basis of assessment becomes all the more difficult the more individuals seek to avoid or evade paying their due." (Hirsch, 1976, 148)

However, just like as *if* morality, as *if* internal motivation is not the same as the real thing, which appears to make it vulnerable to scrutiny by the individualistic calculus. This means that for as *if* internal motivation to function, we need to believe that it is internal and not merely internalised. As soon as we can find that we can abandon it to ensure we will not lose out in cooperative schemes, it is rational to do so.

Considering the inability of external motivation in any form to withstand the force of the collective action problem, we would expect cooperation in society to steadily decline and eventually vanish. But strangely enough, the opposite seems to be the case: since the publishing of the first edition of Hirsch's *Social Limits to Growth* in 1976, the world economy has only grown more, social norms are still in place and no noteworthy riots due to inequality, as predicted by Hirsch (Hirsch, 1976, 149) have taken place¹⁶. On the contrary, there seems to be more collective action than ever before: new institutions and unions are formed daily, even in the secularised West, which should have lost the remainder of its *as if* morality by now. The apparent loss of moral consciousness as the foundation of laws has not led to their abandonment. Instead, many new, contemporary laws have been enacted to address the problems of our time, for example, cybercrime or pollution. It is hard to imagine such laws being proposed by agents with only individual benefit in mind, which gives way to the idea that some form of moral thought is still present, even though this would be impossible given that it should be unnatural and could no longer be supported by external motivations. Even the supposed dichotomy of economy and morality has not resulted in the abandonment of moral discussion on economic subjects, as attested by recent ideas on

¹⁵ We could describe this as *as if* internal motivation

¹⁶ This does not mean, however, that our current economic and social systems are untroubled. Economic and social inequality is still growing on a worldwide scale, giving more cause for riots. But as riots also need cooperation to be effective, it may be so that they are limited because of apparent declining cooperativity. If this is the case, increasing individualisation may be the cause of inequality as 'losing' individuals find it harder to unionise and form powerful associations.

noxious markets by Debra Satz¹⁷ or the discussion on ethical limits to markets prompted by Elizabeth Anderson¹⁸, among many others.

But if external motivation cannot support morality or cooperation as a whole, as we already predicted when reviewing Olson's solution to his collective action problem, then how come both continue to persist? Could it be that there is indeed a natural, internal motivation that supports these schemes of cooperation and morality? And if there is, how does it stand with our rational self-interest that we have henceforth also taken to be a natural or default state of our being? And most importantly, will the faculty of such an internal motivation prove to be a proper solution to the collective action problem? We will examine these questions in the next step in our dialectical inquiry, in which we will analyse the deontological argument for morality made by Immanuel Kant and contemporary theories concerning deontological collective agency.

¹⁷ Satz, Debra. "Noxious Markets". In *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*, 92-113. Oxford: Oxford University Press, 2010.

¹⁸ Anderson, Elizabeth. "The Ethical Limitations of the Market". In *Value in Ethics and Economics*, 141-167. Cambridge: Harvard University Press, 1993.

3. Collective Action As Moral Necessity: Deontology As A Possible Solution

3.1 The Principle of *Ought Implies Can*

We have found in the last chapter that as *if* morality and internalised external motivation do not appear to be stable grounds for perpetual collective action. Rational self-interest seems to erode institutions of collective action, leading to less external motivation to restrain egoistical behaviour, which in turn makes it rational to further abandon these institutions. However, we also noticed that collective action has not disappeared over the last century. Many forms of collective action seem to thrive. In this chapter, we will look into the possibility of internal motivation for collective action. As we have seen in the latter chapters, external motivation means that individuals experience a pressure to do something outside of their own means or desires, like a looming punishment. Internal motivation is signified by internal pressure to do something, either because it is in my nature to do so or because I have a certain internal rationale to do so.

Immanuel Kant has famously argued that rational beings can find morality through their rationality (Moran, 2022, 7). He holds that a moral law, if it were to exist, should be both necessary and universal (Kant, 2011, 84, 86, 88). What he means by this is that any rational being should (necessity) find the same (universality) morality through a process of reasoning. As beings without rationality could never find this moral law, they are not subject to it (Kant, 2011, 98). Adversely, if a being is rational, it means it is, by definition, also morally inclined. Moral duties in deontological theories such as Kant's are thus bound to rational possibilities: if a moral duty is rationally impossible, it cannot be a moral duty at all. This means that irrational beings can never have a moral duty, but it also means that duties that appear moral but imply physically or mentally impossible situations are not moral duties. This notion is known as *ought implies can* (which we will refer to as OiC henceforth): *if* a rational being *ought* to act morally, it has *to be able* to do so.

3.2 *Ought Implies Can* in Collective Action

But there are plenty of cases in which either we are not physically capable of acting morally, or in which the desired ends of acting morally cannot be reached by our own moral decisions. Imagine we encounter a house fire that can only be put out if multiple people form a line and pass buckets of water to one another. Now, it may be the case that I have lost my arms in an accident. As such, I am physically incapable of passing buckets around. Given OiC, my physical state should mean that I am not obligated to help. But when we give an example of impossibility of actualisation, the situation becomes tricky: given that one individual could never put out the fire, it appears that it is impossible to reach the desired moral end from the individual perspective, and as such, no individual has an obligation to put out the fire. If this is indeed the moral situation that deontology supports, it appears that here, too, the collective action problem persists, albeit in a different way: it would mean that no moral duty could exist for actions that need collective action to succeed. Collective action as such remains completely optional, and as we have found before, there appears to be little reason for a rationally self-interested individual to voluntarily engage in it.

Contemporary deontologists have identified this issue and have attempted to resolve it by using the notion of 'group agency', the idea that either an individual's moral duty may extend to the creation of groups that are capable of collective action problems that single individuals cannot, like putting out house fires, or that 'moral groups' readily exist and exert moral duties onto their members. Stephanie Collins argues that for a group to perform moral actions, it must be both an agent and capable of moral understanding (Collins, 2016, 231-232). For a group to be an agent, it must be able to act, that is, it must be able to move its members towards a certain aim. To do so, it needs to be able to have internal discussions on what is to be accomplished, it needs to be able to establish aims based on these deliberations, and it must be able to assign roles to its members to achieve its aims. (Collins, 2016, 232). Collins calls groups that meet these conditions 'collectives'. Furthermore, a collective is a *moral* agent if its aims stem from moral deliberation (Collins, 2016, 232). Collins then argues that even though individual agents have no moral duty to help in situations that cannot be resolved by them alone, they have a duty of responsivity to create moral collectives in emergencies (Collins, 2016, 241). Let us return to the case of the house fire. Even if I may not have a duty to help carrying buckets around, I can see that forming a collective may create an agent that has the capability of putting out the fire by assigning tasks to its members. Because I have a duty to be responsive to the situation and I find that it can only be resolved by a collective, I should collectivise. Then, my duty to further the goals of the collective through fulfilling the tasks assigned to me would serve as a means of indirect duty to assist in situations that I individually would not be able to resolve.

Here, Collins notes that a group is a collective when its members have a shared decision-making process. She posits that for such a process to originate, each member needs to be able to propose group aims, propose roles to achieve those aims, propose a division of the roles, and abide by all aims, roles, and division of roles that is ultimately decided upon. Additionally, members must believe that other members will do the same, and members must have enough knowledge of the proposed aims, roles, and division of roles to abide by them (Collins, 2016, 236).

Nevertheless, there are still two concerns we can have about the duty of responsivity to form a duty to collectivise. Firstly, we may ask whether an individual can more or less correctly assess a collective's chances of succeeding before joining. If an individual agent cannot know whether a collective will succeed, OiC still indirectly applies: I do not know if the collective can resolve the problem, so I cannot have a duty to be responsive and join it. The second concern directly relates to the first as well as to our more general concerns with collective action so far: the need for a belief in other members to abide by the proposed aims, roles, and division of roles that are set by the collective. I can only believe the collective to be successful if I believe that its members will contribute to the cause. However, since they face the same situation, it seems plausible that I cannot reasonably assume they will perform their tasks, let alone assume that they perform them well enough that the collective aims will be met. Since I have reason to believe that the collective aims will not be met if others do not participate and I have equal reason to believe that others will not participate because they cannot be sure of participation of others in the same way I have my doubts, it seems reasonable for me to believe that the collective will not be successful. It seems as if my duty to collectivise is lifted.

Collins addresses the first concern by introducing the concept of probability of success. As we have already seen, if there is no chance of success ($p=0$), OiC dissolves the duty to act. However, it is seldom the case that there is either a zero percent or a hundred percent chance that an act will succeed. Collins thus holds that if the chances of success are very small, especially when the individual costs of acting are very high¹⁹, an action cannot be morally obligated. However, the opposite also rings true: if there is at least some chance of success without too much risk for the self-preservation of the agent, an act may be morally obligated. Collins argues that the same should be true for the duty of responsivity: if I have reason to believe that a collective moral aim can be met successfully if I am to be responsive and collectivise²⁰, I am morally obligated to do so (Collins, 2016, 241). Additionally, Collins believes that once I have recognised the likelihood of collective success, my duty becomes threefold: firstly, I have a duty to collectivise. As soon as I have collectivised, two more duties arise: to fulfil my specific role to maximise the likelihood of success and to maximise the likelihood of collective success in general. Collins states that the latter duty involves convincing other agents of the likelihood of collective success and persuading them to collectivise as well (Collins, 2016, 244-245).

However, it seems implausible that these arguments can overcome our earlier-mentioned concerns about the duty of collectivisation through responsivity. For one, it seems all but proven that we can have accurate knowledge of the likelihood of success of a collective, especially our duty seems to be necessarily tied to that of others (if either one of us does not pass buckets around, there will be no chance to put the fire out, ergo, *not can implies not ought* and both our duties are lifted). Especially with greater collective action problems, like trying to keep climate change under control, the success of the collective depends on so many individuals that it becomes seemingly impossible for each of them to assess whether enough support can be gathered and the problem can be tackled successfully. We can even see an erosive working as we have also seen with external motivation: if I feel like the problem cannot be overcome, whether I collectivise or not, because too little is done by other individuals, my defecting can lead others to do the same, *et cetera*.

3.3 Plausibility of *Ought Implies Can*

There is another problem. Let us take for granted that we as individuals have deemed the chances of success of a collective to be likely and its goals to be morally just and have thus collectivised. Collins argues that as members of this collective, we would now individually have the same duty as the collective. Torbjörn Tännsjö has argued that this reasoning appears shaky: even when collectivised, our individual duties remain contingent on others doing theirs, and as such, our duties are voided if we reasonably assess our efforts to be

¹⁹ It may seem like Collins is stepping out of the deontological realm here. But this is not the case: consider for example a man who might save a child from falling out of a tree ($p=0,001$), but he has a considerable chance of getting himself killed by falling out of the tree himself ($p=0,1$). When asking if it is permissible to not try and save the child in order to not risk death, it seems to be permissible because we need not only consider others as means instead of ends, we must do the same to ourselves (Moran, 2022, 16). Sacrificing ourselves for a small chance of saving someone else would thus not be morally obligated.

²⁰ Given that collectivisation is mostly likely to result in success. If the chances of success are higher if we do not respond by collectivisation but by taking immediate action (because we see others taking action, or not taking action), the duty of responsivity remains an individual duty.

futile in the face of others' actions or inactions (Tännsjö, 1989, 223). As a result, when a collective fails because some or all of its members did not live up to their tasks, this does not result in individual moral failure for any of them: as no one *could* have single-handedly avoided collective failure, no one *ought* to have avoided collective failure. Tännsjö continues to argue that this does not mean that collective duty does not exist, as we can still identify collective success and failure (Tännsjö, 1989, 223-224, 228) and attribute responsibility for it to a collective (Tännsjö, 1989, 228). However, if OiC holds, we are seemingly still left with a void of responsibility between collective duty and individual duty.

But does OiC hold? Many have argued that it does not, at least not in the sense needed to have moral implications. Walter Sinnott-Armstrong has argued that the relation between *can* and *ought* is not entailment, but rather only has conversational implications (Sinnott-Armstrong, 1984, 249-254). This means that even though an action may be or may have become impossible, the duty may still exist or persist. He argues that entailment would make it so that an agent's duties would be lifted if they would become unable to fulfil them, even if they were responsible for this inability (Sinnott-Armstrong, 1984, 252). Sinnott-Armstrong argues that this cannot plausibly be the case, since it would void the moral meaning of duty. Moreover, in case of promises, a duty may be created that could be unfulfillable from the start. If the promise ought not be fulfilled because it cannot be fulfilled, promises similarly lose their moral meaning (Sinnott-Armstrong, 1984, 253).

Sinnott-Armstrong then posits that the only way in which OiC could hold would be on a conversational level, either to advise one to act a certain way or to blame one for not fulfilling a certain duty (Sinnott-Armstrong, 1984, 255-259). For example, if I promised to go to the movies with David, but the public transport is delayed, so I will not be able to make it on time, and I call Mary and ask her what to do, she may respond that I ought to keep my promise. This might be either because she does not know that I am physically unable to keep my promise and may thus serve as a piece of advice, or she knows and she implies that I have not taken appropriate measures to keep my promise, for example, by calling a taxi. In either case, my duty to keep my promise does not dissolve upon my inability to do so.

However, if OiC only holds pragmatically instead of semantically, the moral considerations of deontology become shaky as well: if OiC has no logical necessity or universality, it seems impossible to claim it to be a moral law that can be derived from internal reason, as duty becomes contingent on idiosyncrasies (Mizrahi, 2009, 31).

Although counter arguments against the position of *ought does not imply can*, and *ought only conversationally implies can* have been made by deontologists like Frances Howard-Snyder²¹, whether OiC holds seems to matter very little with regards to deontology's ability to solve the collection action problem: either the principle does not hold and the deontological moral framework as a whole becomes shaky, or the principle does hold and deontology cannot supply a functional moral framework for collective action due to its constraints on individual ability with regards to collective action.

²¹ Howard-Snyder, Frances. "‘Cannot’ Implies ‘Not Ought’", *Philosophical Studies* 130, no. 2 (August 2006): 233-246.

Nevertheless, we may have found a lead to solving the problem. Recall that Tännsjö posits that even though collective duty does not imply individual duty, he still considers the former to exist meaningfully. This would mean that collectives can come to be without collectivised individual duties. This would be strange in both positions we have discussed: the essentialist would have no problem explaining how collectives are either naturally or divinely given, but would then run into the problem of defining what its natural or divine duties are and how individuals are bound by them. Deontology, on the other hand, cannot explain how collectives can have duties if they are not themselves rational agents. It seems impossible to explain what collective failure looks like if duties do not trace back to individuals. An explanation for Tännsjö's position that we have not yet explored would be to turn around the notion of duty and the notion of responsibility. In deontology, responsibility seems to follow from duty: if you ought to do X²² and you do not, you are responsible for not doing X, and as such for the consequences of not doing X. In the next chapter, we will discuss the existentialist notion of absolute responsibility and what it would mean for duty and collective action.

²² And have the ability to do so because of OiC

4. (Inter)personal Existence: Limits to Existentialism

4.1 Existentialism: Existence Precedes Essence

After analysing deontology and (religious) essentialism, there seems to be a clear parallel between the two: both ascribe some form of inherent need for cooperation to humans. In essentialism, this need is seen as a natural or divine disposition, whereas in deontology, it is taken to be a necessary logical²³ condition for being human²⁴. Both rely on the idea that the need for cooperation, or morality, is something sub-human or super-human and should thus form a robust foundation for humans when they are interacting in a subjective and ever-changing world. However, both ideologies face concerns with regard to this robustness.

Essentialist morality appears to need a subjective interpretation of objective morality to codify it and to have real-world implications for individuals' actions. As morality is effectively a handicap when we take for granted that all people are rational egoists, these subjective interpretations weaken it since individuals may claim that their interpretations of it may be equally valid, thereby denying any interpretation of morality that they may consider detrimental to their personal needs and wants. Moreover, this will eventually void the meaning of natural or divine morality, as it can be interpreted in any way.

Deontological morality relies heavily on the concept of OiC. If this notion were to be false, at least semantically, morality would lose the conditions of necessity and universality, effectively voiding its meaning. But even if OiC holds, deontology cannot adequately explain group morality and collective action. Additionally, it is unable to quantify the individual need for cooperation or morality when it comes to collective action, as individual morality cannot depend on the moral considerations of others. As a result, deontological morality seems to rely on the discipline of individuals to watch over their own moral affairs, in which their actions (and their actions alone) are ultimately decisive.

From these concerns, we can derive that a robust solution to the collective action problem not only needs logical consistency but also needs real-world applicability. If we individually find that we *ought* to keep others' needs in mind, but the extent of this duty, especially with regards to collective action problems, remains undetermined, we seem to be left with the same problem in both ideologies, namely that we must all individually determine the extent of our duty towards such problems for ourselves. Given that we are all rationally self-interested, we would then minimise our duty since doing more without guarantees from others would likely make us worse off.

The existentialist Jean-Paul Sartre also noticed this concern. He classifies deontology as a secular form of essentialism, which tries to find the tenets of human nature in logic instead of divine providence (Sartre, 2007, 21-22). He finds that the problem ensues from the desire for an objective human nature, which, after it is found, would be able to give us certain truths in a furthermore rather subjective and ever-changing world. He suggests that we have little reason to believe such an objective human nature exists, but rather that what we see as

²³ Or semantical

²⁴ Or, less specifically, all rational beings

human nature is the culmination of subjective human choices and actions and their worldly consequences.

This is what the ideology of existentialism entails: *existence* (of humans and their actions) *precedes essence* (what is taken to be human behaviour) (Sartre, 2007, 20). The most important difference between existentialism and essentialism when it comes to human nature is thus found in the concept of robustness: whereas essentialists believes that humans are born with a certain, robust internal logic, existentialists hold that what is seen as human nature appears to be immutable, but is determined by all individual human actions and choices, making it contingent (Sartre, 2007, 42-43).

However, it is important to note that this is *ultima facie* precedence: before human nature, there were humans and human (inter)actions, but in the world, humans may base their actions and decisions on their experiences with other humans or humanity in general. Nevertheless, the notion of *ultima facie* precedence brings about a crucial difference between existentialism and essentialism: the notion of absolute responsibility.

Given that there is *ultima facie* precedence, it means that every human agent is born without essence. The way they see themselves and the world around them is formed by their own experiences, actions, and choices. As such, they are ultimately responsible for every action and choice, as any action or choice can ultimately be traced back to earlier actions and choices that an agent has made of their own accord. This responsibility also stretches out to their contribution to humanity or 'human nature': their choices and actions contribute to the existence of humanity as a whole, so they are responsible for aligning their actions and choices with what they want humanity to be (Sartre, 2007, 23-25). Here, we can see a parallel between existentialism and deontology: both believe that we must act in a way that we could see our actions universalised. The difference is that deontologists believe that this follows from reason itself, whereas existentialists believe that this follows from the notion of absolute responsibility: we act in the world, so we shape it and by shaping it, we influence others to act in the same way we do, which means that if what we do would be counterproductive to our existence if it would become common practice, we act in bad faith²⁵ (Sartre, 2007, 25, 43, 46).

4.2 Existentialist Morality

What signifies the difference between Kant and Sartre is the scope of responsibility concerning morality: Kant believes that morality exists in itself, even without moral beings. It is the responsibility of moral beings to find, accept, and adhere to this moral reality. Existentialism broadens the scope of responsibility to a point where morality itself has to be constructed, and humans are responsible for its composition. Morality then follows as a human condition shaped through the actions of responsible individuals. As such, for an existentialist, any action is a moral action. The human realm *is* the moral realm, instead of it existing independently (Sartre, 2007, 33, 36-39, 42-44).

As the human realm and the moral realm are the same, this means that anything and anyone human that comes into existence also exists morally. However, the contents of this

²⁵ That is to say, we act as if we do not think others may act as we would.

morality are still decided by human actions and interactions, which makes morality a part of the human *condition* instead of a human *nature* (Sartre, 2007, 42), as the meaning of morality is what we make it to be by our actions and interactions.

By holding that there is no moral meaning beyond actualisation, existentialism seems to bypass the problem of OiC. Where Kantians ascribe morality only to actions of which the goals have at least some chance of being fulfilled, existentialism poses that any intention has no moral meaning until it is acted upon and that no moral fulfilment can be permanent, as the human condition can never reach a state of completion (Sartre, 2007, 52). As moral meaning is in the act itself, no hope of success is required to have a responsibility to act and to bring meaning into the moral (and human) realm (Sartre, 2007, 36-37).

One parallel between Kant and Sartre remains, however; although they differ on whether being moral is itself primordial or follows from a primordial rationality, they both take it to be a necessary condition of being human. Sartre notes that humans cannot meaningfully exist as beings-in-themselves: they have a place as beings-in-the-world and can only define themselves and the world by relating to the world (Sartre, 2007, 52-53).

4.3 Problems of Existentialist Morality

Where does the existentialist stand in relation to our collective action problem? Both Sartre and Olson seem to believe that humans are, *ultima facie*, free, which means that their actions are never necessarily bound to any nature or social construct. However, their assessment of the human condition differs strongly: Olson acknowledges that there are better outcomes than those usually put out by Prisoner's Dilemmas, but that those are very hard to attain if the outcomes depend on the actions of others. Existentialists, on the other hand, would claim that because we always depend on others for outcomes, we cannot forfeit our responsibility to make our own moral choices or let the consideration of outcomes alter our choices. There are no excuses for any moral choice, because we are fully responsible for personal action (Sartre, 2007, 29).

However, the strong notion of responsibility that existentialism posits makes for moral situations that may seem non-intuitive or even implausible. The notion that the human condition, and thus morality, is ever-changing means there is only moral meaning to moral means and none to moral ends. But when we are acting morally, we do so with certain ends in mind. Even if these are purely the ends that we individually find worthy of dedicating our actions to, it means that we still have to ascribe some meaning or value to them. If we take any end to be defined by incompleteness and infinity, the term becomes a logical contradiction. We may well call this the *problem of the final ordeal*: if no meaning can be attributed to any end, that would also include the meaning of our actions concerning that end. In the infinitely great scheme of things, any action could lead to any possible world in which that action may or may no longer be possible, or may or may not become common practice. Sartre holds that no hope of success is required to act, but what if 'success' as a concept becomes impossible, and no action seems to have any direction?

In a sense, it appears impossible that in such a world, any consideration of morality we may have through experiences and earlier actions has any impact on the morality of our

immanent actions. This seems to leave our actions with an almost Schrödinger²⁶ morality: we cannot possibly foresee if we would deem our actions to be morally good or morally bad in the greater scheme of things, nor can we assess actions of others in this regard. This means that even though all action is moral, the contents of this morality are so subjective that it appears incredibly hard to appeal to them on a retrospective level, let alone in a social context.

This also leads us to a second argument of counter-intuition that can be made against the existentialist position, an argument of fairness. If the morality of actions is indeed so subjective that we cannot recognise the *ultima facie* moral impact of our actions before acting, it seems unfair to be held responsible for this impact. Even if we are absolutely responsible by necessity, to be (fairly) held responsible is commonly taken to also be an important part of real-world morality. It feels unfair to hold someone responsible for the moral consequences of their actions if there is no guideline at all.

The existentialist may reply to these concerns by proposing that even though there is no ultimate end to the human condition, it is nonetheless at all times and our continuous interpretation of it should give us the desired guideline on what is considered morally good or bad *at the time*: considering all experiences and actions I have undertaken in the human condition, I can confidently say that *at this time*, it is wrong to kill another human being. Of course, the wrongness or rightness of not killing another person may change over time, and we would still be responsible for our actions, but as non-actualised situations have no moral meaning, we should not let them be part of our moral concerns.

However, one of the problems we identified in deontology seems to concern the existentialist morality also: it mostly deals with individual morality through individual actions. Any individual has their own experiences and takes their own actions, which makes it so that they have their own interpretation of the human condition. Even though different individuals have some form of necessary overlap insofar as they both exist in the same world and they can only meaningfully exist through their relationship with the world, these relationships themselves are uniquely tied to the individual. As such, we may ask if there is any overarching concept of the human condition or if it is purely an individual interpretation of what it is, and ought to be, to be human. In the latter case, individuals may have vastly different views of what morality or cooperation should look like, giving us no insight into why cooperation often succeeds, even with Olson's rationale present.

But despite this, we may have found a very constructive piece of the puzzle already while we approach the final chapter. Contrary to essentialism and deontology, existentialism seems to successfully create a robust, ground-level concept of morality and cooperation: morality *is* the way we interact with the world, the way we give it meaning, and with it, give ourselves meaning. The same appears to ring true for cooperation: we find meaning in our individual world by relating our experiences, feelings, and actions to the people around us. This meaning can only be relationally explained, which means that the starting point of cooperation is to acknowledge both the existence of other people and the necessity of

²⁶ Erwin Schrödinger (1887-1961) was a quantum physicist who became famous thanks to his theory of radioactive material being decayed and non-decayed at the same time as long as it is not being witnessed.

relationality for our meaningful existence. With these notions in mind, we will proceed to the final chapter, in which we will argue that they are fruitful premises to further explain the success or failure of collective action.

5. *Homo Cooperatis*: the Existentialism and Essentialism of Being Human Revisited

5.1 A Return to Essentialism?

Throughout the last four chapters, we have examined different approaches to a two-sided problem of collective action by means of regression and erosion. First, we analysed the contractarian and contractualist arguments for collective action and concluded that they attempt to solve the regression problem by introducing a divine or natural starting point to collective action. With such a divine or natural starting point to fall back on, the problem of erosion should also be cleared. However, we found that for such a starting point to work, its contents should be objective and clearly definable. If the natural or divine starting point would be dependent on individual interpretation and acceptance, a situation will ensue in which every individual will choose the interpretation that yields the most individual benefits, which in turn leads to erosion of collective action as first free-riding becomes a viable strategy to increase benefit, eventually leading to abandonment of cooperation as a viable beneficial strategy altogether as no individual can be trusted to cooperate.

The same problem occurs in the essentialist reasoning: without objective moral truths, there is no reason to believe others will keep their promises, especially if breaking these promises would lead to individual economic benefits. A possible solution, as *if* morality or internalised morality, does not seem to be able to defend against the erosion problem, as internalisation is not the same as true internal (intrinsic) reason: it still depends on external pressure to uphold its premises, making it vulnerable to abandonment when the external pressure diminishes and the advantages of cooperation are outweighed by the advantages of free-riding.

The deontology we discussed in the third chapter thus focuses on finding intrinsic reasons for cooperation in the form of independent moral laws that are universal and omnipresent, even if moral beings did not exist. For this reasoning to work, moral beings must be able to find the premises of such laws by their own accord, and they can only yield duties that can be successfully fulfilled. This principle of *ought implies can* appears to be problematic in twofold ways: firstly, if it were to be true, it seems that there can only be individual duties, which in turn would mean that seemingly moral problems which cannot be solved by single individuals are outside the scope of individual duties, leaving them in moral limbo. Secondly, OiC appears to be shaky ground for individual duty as well, as it seems counterintuitive that moral duties can easily be evaded by rendering their fulfilment impossible, for example, when we make a promise that we know we cannot possibly keep. Deontologists have tried to solve the first problem by appealing to a duty of collectivisation, but this duty also hinges on the individual assessment of the collectivisation process and the possible collective having possibility of success. As the individual cannot assume a positive assessment of these processes in other moral agents, it cannot come to a positive assessment itself, effectively clearing its duty to collectivise.

Sartre's existentialism tries to counter these problems by stepping away from a pre-existing morality, human nature, or divinity. He argues that morality is a necessary part of the human condition, which he describes as the state in which humans find themselves after birth and

which is indefinitely shaped by individual experiences and actions. The contents of this morality are then formed by the individual's need to act because one is ultimately responsible for every action (and inaction), making for passive collective action through a totality of individual actions. However, as this passive collective action does not refer to a singular human nature or even to a singular human condition, it appears to lack universality, making the contents of this passive collective action seem to appear by chance. It also gives little insight into the necessity of active collective action and the possible contents of such collective action.

It appears our current version of existentialism and its suggested relationality between individuals may be too focused on the individual freedom to decide the contents of their perceived human condition, as well as morality, to explain active collective action, and to suggest what its contents could or should be. What we mean by this is that the individual appears to fully decide how they view themselves and others around them: an individual is at the basis only responsible for itself. Its individual existence still precedes its morality, effectively making it so that essence and the existence of others follow from their perception.

A notable critique on this view comes from Emmanuel Levinas: he suggests that other individuals are not merely perceived and defined from the position of existence of an individual, but that they can never be fully grasped as their existence has its own individuality, with an inner world and experiences that the initial individual can never comprehend (Levinas, 1969, 39). He then proceeds to say that our relation with others thus has a special, conversational aspect: in order to converse with a creature we can never fully understand, we have to question our own understanding of things to try and understand them (Levinas, 1969, 40, 43). In other words, conversing with others makes it so that we actively adapt our understanding and views to those of others, by relating our responsibility to others. Given the existentialist ideals, it is clear why we would do so: without them, we would be without meaning as our actions become meaningful in relation to the world. It also becomes clear that there is a distinction between *conversing about* and *conversing with*, in which the latter takes on a position of precedence over the former: defining the world can only be a meaningful exercise if we converse with a creature that, at least partially, understands our definitions and who is willing to alter, deny or affirm them. This means that meaningful definition requires communication between two like-minded yet unique beings that understand that they can never truly define each other because their ability to define stems from having a limited grasp on the other and the indefinite need for communication (Levinas, 1969, 66).

Levinas then proceeds to posit that this understanding of the other as a limiting yet necessary factor for meaningful definition of the world and, with it, meaningful existence in the world, leads to ethics, which he defines as the process of questioning oneself and justifying oneself towards another (Levinas, 1969, 40, 43). In this reasoning, we can define ethics as the contents of morality, which in turn is more elaborative than the existentialist version we have encountered before: in addition to morality being an existential human need to act in the world, this acting can only meaningfully take place as, or after, *interaction*. This paves the way for active collective action as a necessary condition for meaningful individual existence.

Interestingly enough, we could argue that the existentialist principle of morality, and especially Levinas's version, implies some form of essentialism: physical human existence and moral essence through necessary action or interaction seem to occur simultaneously, insofar as existence has no true point in which it is without morality: even before self-consciousness develops, action and interaction already take place to fill in the early basis for morality²⁷.

5.2 The Ethics of Hybrid Existentialism

An argument could be made that even though the simultaneous immediacy of existence and essence appears to imply the innate possibility of cooperation, it still gives little reason to believe that cooperation ought to take place. Additionally, an argument can be made that the immediacy of essence makes us unfree, which could lead to the implausibility of accountability. Both concerns are connected as both confuse morality with ethics (or, the contents of morality). The second concern assumes that essential morality implies essential ethics, whereas the first concern assumes that essential morality does not necessarily imply any form of ethics. Both implications cannot be true if existentialism is to provide an answer to our problems of collective action.

If essential morality were not to necessarily imply any ethics at all, it would mean that the essentiality of morality, to create meaning, could simultaneously be dormant and still exist meaningfully. However, we have seen before that any action has moral implications (and thus leads to ethics). The creation of ethics is what logically provides morality with meaning. This means that a dormant essential morality cannot meaningfully exist. As such, essential morality has to imply some form of ethics. If this ethics were essential as well, it would mean that no action or interaction could change its course: the human condition would be indefinitely determined, in other words, we would just be along for the ride. If this were to be true, it would mean that there would be no agency, and actions and interactions would thus only be "events". This seems implausible for similar reasons as pure essentialism seems implausible: if this implication is true, ethics should be identical for everyone, as our interpretations should not be able to change the essentiality. Additionally, it would mean that there would be no need for interaction or action to define the world, as the definitions would be implied. This means that if essential morality would imply essential ethics, it would, in turn, render morality inessential. As this is a logical contradiction as well, it means essential morality cannot imply essential ethics.

What form of ethics *does* our version of existentialism imply, then? As we have found before, morality must create ethics, first specifically by human interaction and later by all human action. For this interaction to take place, both individual agents must assume that they cannot fully know the existence of the other, and they must thus limit their actions in order to create meaningful interactions, as these interactions form the basis for meaningful actions. In cooperation, this would mean that any individual action is only meaningful while it is either part of cooperative action or follows from a certain scheme of cooperative action. To act without taking into account which scheme of collective action the action relates to or is based upon would be to deny an important part of its meaning, namely the part where other

²⁷ Note that these pre-conscious interactions are very similar to the animal instincts we discussed in the first chapter!

humans necessarily co-define its contents. As actions cannot have meaningful content (or ethics) without interactions and essential morality necessarily implies a form of ethics, it means that all ethical content, actions and interactions alike, must be interactive.

5.3 Problems of the Hybrid Existentialist Ethics

What this means in practice is that meaningful individual existence can only persist through fundamental interaction, and individuals who pretend their actions are not tied to interaction undermine their own existence. In other words, not only are individuals essentially interactive, but all their actions must follow from some form of interaction, meaning that they must necessarily limit their actions with regard to another human. But if there is omnipresent interaction, why does collective action still fail sometimes?

We can give three explanations for this phenomenon: firstly, note that, theoretically, ethics only requires one to limit their actions with regard to at least *one* human being. This means that human beings that one does not care to relate to may be deemed irrelevant for one's limitation of actions. The second explanation would be to state that, because passive collective action exists (social consequences of individually motivated actions), agents necessarily meet the moral requirement for meaningful existence, and thus, intentional interactivity would be rendered unnecessary. The last explanation considers the ability of individuals to intentionally and non-intentionally 'overlook' their ethical position by denying (intentionally or non-intentionally) that their individual meaning is tied to other individuals.

None of these three explanations leads to an excuse for unethical or non-cooperative behaviour. The first explanation is flawed because it assumes that individuals can take on the full meaning of the world by relating to a limited number of individuals. Additionally, it assumes that an individual can precisely monitor what amounts and contents of meaning are brought into their life by whom and can actively decide to halt the meaning one brings into their life. It is precisely this lack of control over other individuals that we have found earlier to be crucial for morality to exist.

The second explanation assumes that there is a categorical difference between 'interactions' and 'actions that have interactive consequences' and that the latter category is sufficient for morality to exist. This reasoning would render morality to be a byproduct of individual actions, which would mean that our meaningful existence would be based on chance encounters. Apart from the fact that this seems counterintuitive, it also contradicts logic, as we have found that for morality to yield meaning, actions would have to intentionally be limited by other individuals, which is not the case if meaning is only a byproduct of chance.

To elude this, we can take, for example, the definition of an apple. For the word 'apple' to truly mean something, two agents must agree that the word signifies an object in the world. There would be no reason for an agent to make up the word 'apple' purely for himself: they do not need to communicate with themselves. Then, if they try to set up the word as a medium for communication, they still have to make the other person understand what they means, and the other person can choose not to accept the word or propose another word. In any case, the word can only be meaningful if it is actively made a medium for communication; it has no reason to exist for one person or separate persons without a need for communication.

Now, one may argue that although this means that interaction is necessary, it does not mean that interaction has to be cooperative. Other forms of interaction, such as domination or competition, seem to have little cooperative aspects. However, this only seems to be the case *prima facie*: the *ultima facie* purpose of domination is to make subjected agents cooperative, which fails if these agents do not cooperate by not accepting dominion. And in cases of competition, *ultima facie* cooperative ground rules appear to be necessary on what the competition is about and what methods are allowed. Both domination and competition appear only to exist as interactions concurrent with, or after, cooperative interactions. This adds another piece to our puzzle: the contents of morality we are after are not merely interactive, but cooperatively interactive.

Having shown why the first two possible reasons collective action fails seem to rest upon incomplete understanding of ethics, we move on to the last, and possibly most common, reason for collective failure. Similar but broader than the first two reasons, the last reason considers the human inability to recognise the tie, or at least the extent of this tie, between their existence and the moral principles and ethics we have discussed. It comes down to the question, “Why would we interact cooperatively if we are free not to?”. In other words, what makes an *ought* to a *need* to, while still maintaining individual freedom? It is important to differentiate here between *prima facie* and *ultima facie* freedom: on the surface level, one has complete freedom (and complete responsibility) to act in any way one seems fit. But fit to what? Fit to bring meaning into their life by acting morally, as they are a moral being. They are, *ultima facie*, not free to choose whether or not they are a moral being, as this morality is necessarily tied to the meaning of their existence. And to act by their morality, the only acts that they can deem to be ‘fit’ are acts that have a basis in cooperative interaction. As such, any act that denies or even deteriorates this basis is an act that denies an essential part of its meaning. Such acts are logically contradictory and cannot be excused by means of freedom or ignorance.

5.4 Implications and Considerations Concerning Hybrid Existentialism

Now that we have defined the reasons for failure of collective action as non-rational, we can shortly analyse why they are still used and how we may avoid their use. Let us first reminisce about Gauthier’s premise from the first chapter about conditional cooperation. As we are unable to tell whether others will cooperate without former cooperation, the reasonable default state of individuals appeared to be conditionally cooperative, but non-cooperative when actualised. Similarly, we can make the argument that ignorance of the necessity of morality for one’s meaningful existence can lead to this situation: if I take a meaningful individual existence for granted and only see others as agents that have no initial connection with me, I have little reason to engage with them other than when I can be sure they will add something to my existence. According to our current findings, this thought process is mistaken: other agents do not merely add to my existence; they form the medium in which I exist for myself. As such, there is already a cooperative basis (morality) without their being cooperation (ethics). We could say that humans are conditionally cooperative *prima facie*, but unconditionally cooperative *ultima facie*: they can choose forms of cooperation, but they cannot choose non-cooperation. And in choosing the form of cooperation, they are provided with meaning as well as responsibility for the meaning they created. To make people more cooperative, then, seems to be a matter of raising their

awareness of their situational freedom, their need for cooperation, and their shared predicament.

We have now come to a point in our analysis at which we can give a fairly convincing explanation as to why collective action has not - and will not - come to a halt even though external motivation should regress infinitely. With individuals having an inherent need to cooperate in some manner to exist meaningfully, we have found a form of motivation that cannot be regressed infinitely, thereby ensuring a base level of cooperation in humanity that can always be built upon as long as it is recognised. However, this also signifies three problems that collective action will necessarily face: a problem of recognition, a problem of quantifiability, and a problem of regulation.

Let us first look at the problems of quantifiability and regulation. What we mean by this is that it remains unclear to what extent we must reasonably cooperate and if any forms of cooperation should have priority over other forms of cooperation. These problems are reflected in common beliefs in other philosophical currents that deal with morality, such as care ethics, communitarianism, Adam Smith's mutual sympathy, and Peter Singer's utilitarianism. Singer has famously proposed that our moral decision-making should consider *all* moral predicaments that we can conceive, even those at great mental or physical distances (Singer, 2022, 250). This would, in turn, mean that the extent of our cooperation should be as far as possible. Singer also believes that if we have the ability to prevent or dissolve a bad moral situation without creating a worse moral situation, we have a duty to do so (Singer, 2022, 249). This means that the duty to cooperate would be based on the eventual moral outcome. If we were to take this stance, we would have to settle for an incredibly broad scheme of obligated cooperation, to such an extent that we may ask whether this much cooperation is truly needed for meaningful existence, or, similarly, if this amount of cooperation necessarily follows from our need for cooperation.

It is for this reason that both care ethics and communitarianism have proposed that cooperative duties weigh heavier when individuals have a closer mental or physical relation. Communitarians have suggested that physical proximity, historical proximity and psychological proximity are all important notions to consider whether cooperation should and will take place (Bell, 2020). Care ethicists have endorsed this sentiment, stating

"it is physically impossible to 'care for' all of humanity, strangers who have not addressed us directly, or those unknown others at a great distance" (Gomez et al., 2007, 134)

Both communitarianism and care ethics thus reject the idea of universal moral laws and duties (Gomez et al., 2007, 134 and Sandel, 2012, 1-14) and instead prefer to focus on cooperative duties within a smaller scheme of cooperation. For the purpose of workability, as Olson also suggested (Olson, 1971, 53-57), taking a smaller scheme would probably make cooperation more effective as the benefits of cooperation to oneself and the world would be more visible²⁸. However, using this scheme would mean that we would need to believe that some people are out of our cooperative scope, or at least that cooperative duties are ordered in such a way that we may act as if people at a greater physical or mental distance add less to our meaningful world. This in itself brings back the risk of regression: it invites us

²⁸ One may also recall this from page 7 of this thesis.

to define a point at which we do not have to cooperate anymore and by defining this, we also open up the possibility of this definition shifting to a different position, which could narrow down our cooperative duties to a point at which they essentially become meaningless.

Adam Smith agrees with the communitarians and care ethicists that our morality seems to depend on proximity, as he has famously described in his *Theory of Moral Sentiments*. It is important to note that Smith has tried to describe how the world *is*, rather than what it *ought to be* (Forman-Barzilai, 2005, 191). As such, all morality (or duty) should rest upon existing processes in the world. With that, Smith's project is closely related to our own when we derived morality and ethics from their necessity for meaningful human existence. Interestingly enough, he too seems to acknowledge a Levinasian account of how we come to morality: by *surveilling* the actions of another agent and expressing one's judgement, the other agent may be *disciplined* into altering its behaviour to come to a more collective coexistence (Forman-Barzilai, 2005, 192). Note that this closely resembles Levinas's account of an individual limiting themselves as they try to converse with another agent that they cannot experience as if they were the same agent in an attempt to create a medium of understanding (and meaning) between them. Smith calls this process *sympathy* (Forman-Barzilai, 2005, 192). This process becomes even more Levinasian when we consider the reason an individual would engage in surveillance and judgement: Smith proposes that we do so because of *self-reference*, the idea that we believe the other's reality to be on par with our own (Forman-Barzilai, 2005, 194). This effectively means that we consider the other to be part of our own existence. Sympathy, in turn, can lead to common expectations of whom we should care for and how much. This is what Smith calls *affection* (Forman-Barzilai, 2005, 201), which can be seen as a form of ethics. However, Smith makes it clear that this affection should always follow from sympathy and that sympathy cannot take place if there is too much distance between agents. Moreover, humans are limited in their spectative possibilities and to expect affection for everyone would mean that there should also be sympathy for everyone, which contradicts the natural way in which we have more sympathy for those who are closer to us (Forman-Barzilai, 2005, 201). As such, it is unreasonable to expect affection from everyone, or as Smith states

“All men, even those at the greatest distance, are no doubt entitled to our good wishes, and our good wishes we naturally give them. But if, notwithstanding, they should be unfortunate, to give ourselves any anxiety upon that account, seems to be no part of our duty.” (Smith, 1982, 140)

Now, if we return to the other proposals, Smith seems to offer a solution that seems more intuitive than Singer's cosmopolitanism or the communitarian and care ethicist positions. However, especially since Smith does not explicitly make a normative claim, it remains unclear to what extent we ought to cooperate. In fact, it only becomes clear that the other positions each appear to focus too much on one end of the cooperative spectrum. It also shows that the hybrid existentialism we have found cannot yet adequately deal with questions of quantifiability and regulation.

The other problem we still have to face is that of recognition. We have already found in this chapter that ignorance or wilful denial of the need for cooperation is inexcusable. However, in reality, we find that many forms of cooperation fail for this exact reason, either because agents are not aware of the importance of cooperation in relation to their own meaningful

existence or because they downplay this importance. For example, Singer notes that some moral actions are considered duties, whereas others are categorised as charity (Singer, 2022, 253). By making this distinction, it seems we define some moral situations differently to escape our duty to cooperate. We still acknowledge the good of such a cooperation, but we deny our own duty to contribute to this good. Smith may argue here that the difference is legitimate and relies on the difference between affection and general good wishes. But here too, the question remains if we can make this distinction effectively and in good faith, or if we make such a distinction with the ulterior motive to free-ride. Hybrid existentialism cannot solve this problem at the moment.

5.5 Conclusion

Throughout this thesis, we have analysed several attempts to explain the workings of cooperation with regards to the problem we found in the first chapter. In the second chapter, we noted that cooperation and morality in general are closely tied to each other, as both rely on social dogmas or agreements. Throughout the second and third chapter, we have considered external motivation and internal motivation as possible conditions why people cooperate. We found here that both forms of motivation can explain to an extent why cooperation exists, but have no satisfactory explanation as to why it persists or ought to persist. In the fourth chapter, we have considered the possibility that cooperation (and morality) are the necessary result of our human existence. We noted that for our existence to have meaning, there has to be some form of relationality with other existents. We have elaborated upon this idea in the last chapter, in which we have proposed that cooperation and morality do not merely follow from human existence, but that they are an immanent and necessary part of it. With this theory, we have shown that external and internal motivation may both be part of the solution of the collective action problem: our existence has an intrinsic (or internal) necessity to connect to others (external factor) in order to meaningfully exist, which means that it is impossible not to cooperate with others.

With this in mind, we have considered what this means for cooperation as a whole. We have proposed that “to create meaning by connecting to others” can be defined as morality, whereas the created meaning can be defined as ethics. We have noted that a form of ethics must always follow from morality, but that the contents of ethics remain quite fluid, leading to new problems of collective action in the form of the necessary extent of individual morality with regards to what ethics are produced and ought to be produced. It also brings about the question as to what ethics an individual should adopt and how individuals can be made to recognise their morality and its necessity with regards to their own meaning, especially when many forms of meaning are already out there and are easily taken for granted.

As such, the theory of hybrid existentialism has much room to grow. We can make one proposal as to where to start with the last concern we mentioned: by lifting this predicament of this theory out of the sphere of mutual knowledge and into the sphere of common knowledge, it may be possible to have more overall awareness and a more open debate on the topic of collective action.

What do we mean by that? Mutual knowledge is generally defined as “knowledge that a group of people have, but of which none of the people in that group necessarily know the others have it”. We have already seen that by going through some reasoning, we may come

to the knowledge that we are in a cooperative predicament in which others matter for our own sake, even if they do not know or acknowledge this. However, by informing others of this finding, they may not only find the same predicament via the same reasoning, but they also know at this point that we know and believe the same thing. This may make cooperation easier as well as make it easier to argue about the extent of cooperative duty. By doing so, we may well be on our way to moving from collective action as a problem to collective action as a solution to the real-world problems we are faced with in the current day and age.

Bibliography

- Bell, Daniel. "Communitarianism". At *Stanford Encyclopedia of Philosophy*, pub. 2001, rev. 2020. <https://plato.stanford.edu/entries/communitarianism/>.
- Berlin, Isaiah. "Two Concepts of Liberty". In *Four Essays On Liberty*, 118-172. Oxford: Oxford University Press, 1969.
- Collins, Stephanie. "Collectives' Duties and Collectivization Duties", *Australasian Journal of Philosophy* 91, no. 2 (June 2013): 231-248
- Forman-Barzilai, Fonna. "Sympathy in space(s): Adam Smith on Proximity", *Political Theory* 33, no. 2 (April 2005): 189-217
- Gauthier, David. *Morals by Agreement*. Oxford: Oxford University Press, 1987.
- Gomez, Ana Maria Davila, and Crowther, David. *Ethics, Psyche and Social Responsibility*. Aldershot: Gowel, 2007.
- Hirsch, Fred. *Social Limits to Growth*. Cambridge: Harvard University Press, 1976.
- Hobbes, Thomas. *Leviathan*. Ed. W.G. Pogson Smith. Oxford: Clarendon Press, 1929.
- Kant, Immanuel. "Grundlegung zur Metaphysik der Sitten". In *Groundwork of the Metaphysics of Morals. A German-English edition*, trans. and ed. Mary Gregor and Jens Timmermann, 1-156. Cambridge: Cambridge University Press, 2011.
- Levinas, Emmanuel. *Totality and Infinity: An Essay on Exteriority*. Trans. Alphonso Lingis. Pittsburgh: Duquesne University Press, 1969.
- Locke, John. "A Letter Concerning Toleration." In *Two Treatises of Government and A Letter Concerning Toleration*, ed. Ian Shapiro, 211-256. New Haven: Yale University Press, 2003.
- Locke, John. *An Essay Concerning Human Understanding*. Ed. Jim Manis. Philadelphia: Penn State University Press, 1999.
- Locke, John. "Two Treatises of Government." In *Two Treatises of Government and A Letter Concerning Toleration*, ed. Ian Shapiro, 1-210. New Haven: Yale University Press, 2003.
- Mizrahi, Moti. "'Ought' Does Not Imply 'Can'", *Philosophical Frontiers* 4, issue 1 (2009): 19-35.
- Moran, Kate A. *Kant's Ethics*. Cambridge: Cambridge University Press, 2022.

- Olson, Mancur. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press, 1971.
- Plessner, Helmuth. *Die Stufen des Organischen und der Mensch*. Berlin/New York: De Gruyter, 2019.
- Rawls, John. *A Theory of Justice*. Rev. ed. Cambridge: Harvard University Press, 1999.
- Sandel, Michael J. *Liberalism and the Limits of Justice*. Second Ed. Cambridge: Cambridge University Press, 2012.
- Sartre, Jean-Paul. *Existentialism is a Humanism*. Trans. Carol Macomber, ed. John Kulka. New Haven: Yale University Press, 2007.
- Schaefer, David Lewis. "A Critique of Rawls' Contract Doctrine", *The Review of Metaphysics* 28, no. 1 (September 1974): 89-115.
- Singer, Peter. "Famine, Affluence and Morality", *International Ethics*, Vol. 3 (2022): 247-262.
- Sinnott-Armstrong, Walter. "An Argument for Consequentialism", *Philosophical Perspectives* 6 (Ethics) (1992): 399-421.
- Sinnott-Armstrong, Walter. "'Ought' Conversationally Implies 'Can'", *The Philosophical Review* 93, no. 2 (April 1984): 249-261.
- Smith, Adam. *The Theory of Moral Sentiments*. Ed. D.D. Raphael and A.L. Macfie. Indianapolis: Liberty Press, 1982.
- Tännsjö, Torbjörn. "The Morality of Collective Actions", *The Philosophical Quarterly* 39, no. 155 (April, 1989): 221-228.
- Wilson, James Q. *Political Organizations*. Upd. ed. Princeton: Princeton University Press, 1995.