

September 15, 2019

The Samaritan's Curse **The Challenge of Group Morality**

Kaushik Basu

Department of Economics
Uris Hall
Cornell University
Ithaca, New York 14853

Email: kb40@cornell.edu

Abstract

The paper shows how concepts from game theory can shed light on the long-standing debate in philosophy concerning how we should distribute moral responsibility when a group behaves badly. The concepts make possible new answers to some vexing questions of moral responsibility and the rejection of some of the popular. The paper develops two new games, 'After the Seminary' and 'the Samaritan's Curse', which illustrate the paradoxical result that individuals' becoming moral can lead to immoral behavior on the part of the group. The findings caution us against quick conclusions about individual intent from observations of collective behavior.

Key words: moral responsibility, paradox of morality, Samaritan's Curse, Nash equilibria

JEL Numbers: D62, D64, K13

Acknowledgement This paper has evolved greatly from the time of its first presentation at a seminar in the Philosophy Department of Cornell University on 27 March 2019. I benefited greatly from the comments received during that seminar and later presentations at the conference in honor of Jorgen Weibull at the Stockholm School of Economics, 14 June 2019, and a workshop on game theory at IIT Bombay, 3 August 2019. For valuable suggestions I would like to record my thanks to Abhijit Banerjee, Alaka Basu, Karna Basu, Tad Brennan, Stephen Coate, Avinash Dixit, Meir Friedenber, Nicole Hassoun, Rachana Kamtekar, Theodore Korzukhin, Ajit Mishra, Ariel Rubinstein, Sudipta Sarangi, Eric Van Damme, and especially, Larry Blume and Karla Hoff.

The Samaritan's Curse: The Challenge of Group Morality

1. Introduction

It is not uncommon for us to attribute moral qualities to groups. Here are Nancy Altman and Linda Benesch in the **Huffington Post**¹ discussing the immorality of “today’s Republican leaders” in the US: “The Republican elite’s immorality goes well beyond Donald Trump ... The immorality and disdain of today’s Republican elites shine through in the policies that they embrace.” To ward off any charge of partisanship, let me cite Dave King about the “immoral Democrat Party,”²: “Here is a short list of the damage being done to our country by the immoral Democrat party.” I may just add that his list is not very short.

Such indictments are not exclusive to America. The Archbishop of Canterbury famously lamented that a no-deal Brexit will be a collective “moral failure”. In similar vein, we talk about the lack of morals of big corporations; we use collective moral epithets to describe the Houthis, Tutsis, FARC rebels, and protest groups as in the Arab Spring, Hong Kong’s Umbrella Revolution, and the North Korean leadership. Even as I began working on this paper, there was breaking news in India raising questions of collective responsibility, with huge real-life implications. This concerned the ‘Pulwama attack’ on 14 February 2019, when a terrorist group operating out of Pakistan attacked an Indian convoy. It triggered an emotional dispute about who bore moral responsibility for this act. I shall return to this later.

The question of collective moral responsibility and group agency has been a subject of considerable discussion and debate in law and economics, and more so in philosophy³. How these debates are resolved can make a difference to how we argue about group responsibility, and that can make a difference to how we behave. The glib uses of moral categories in public discourse feed into our thinking about politics, the economy and society, and our own decisions, such as what we buy, what we wear, how we vote, and on some occasions, whether we go to war.

The aim of this paper is to clarify some of these terms and attributions. I do this by bringing in ideas from game theory, developing two new games— ‘After the Seminary’ and ‘the Samaritan’s Curse’, and creating new conceptual categories. The use of game theory in philosophy brings a surprising sharpness to some of the debates (just as it did a few decades ago with Lewis’s work on conventions). When it comes to rational behavior, thanks to game theory, we have reached much greater clarity on attributions of group behavior to individual choice. This has had an impact on public discourse and policymaking. Seeing a group of villagers over-grazing and ruining the commons, we do not lament that the villagers are not acting in their self-interest. Instead, we realize that their self-interest is the source of the problem. The Prisoner’s Dilemma is the classic game that illustrates this. It has helped in the design of policies for solving the commons problem

¹See https://www.huffingtonpost.com/entry/the-deeply-immoral-values-of-todays-republican-leaders_us_5a2eb9f7e4b04cb297c2aee5

² See <https://www.conservativedailynews.com/2015/11/the-immoral-democrat-party-list-of-shame/>

³ See, for instance, Feinberg (1968), Sartorio (2004), Tannsjo (2007), Petersson (2008), List and Petit (2011), Hess (2014), Arruda (2017).

and of agreements among nations and their leaders to tackle global climate change. We are today aware of many of the pitfalls of individual rationality for collective well-being. We have studied teams and organizations made up of rational individuals, collectively making irrational choices that reveal intransitive preferences (see Marshak and Radner, 1972; Rubinstein, 1998, Chapter 6). A related problem arises in what is referred to as the ‘common agency’ problem, which deals with situations where several principals, who have potentially conflicting aims, have to operate through a common agent (see Bernheim and Whinston, 1986).

There is no such clarity when it comes to individual *moral* behavior and collective outcomes⁴. This is the source of many of the conflicts in the world. Once you become sensitized to this problem, it becomes difficult not to be dismayed by the popular discourses on group behavior and attributions of responsibility. This lack of clarity feeds into poor decisions and policies.

This paper shows that not only is it difficult to allocate to group members moral responsibility for some bad actions by a group, as pointed out in a large literature (for recent research, see Bjornesson, 2014; Chant, 2015; and Arruda, 2017), but that at times *individual* morality is the source of *group* immorality. This raises a host of questions about the place of consequentialism and deontological ethics. The paper ends up taking a broad position of normative methodological individualism. It argues that this is the only reasonable way to approach groups and collectivities. It is incorrect to treat a group or a collective as “a *moral* agent in its own right,” with its own “rational point of view” (see, e.g., Hess, 2014, p. 203) and with its “collective intention”⁵. At the same time, as explained in the closing section, it is important to keep alive the ‘moral intention’ with which human beings are naturally endowed and use it to think beyond the game.

Dworkin (1986) was an early exponent of the moral agency of a corporation. In his view, a corporation should be held responsible for certain wrongs as a first step before allocating the responsibility to the individual decision-makers in the corporation⁶. This however may not be easy or even always possible. The Samaritan’s Curse illustrates a situation in which it is not possible to reduce the responsibility of the group to the individuals who constitute it.

The current literature distinguishes group moral responsibility that can and cannot be distributed among the individuals of the group. When it cannot be distributed, “groups sometimes might be regarded as independent agents... Hence, it seems possible that groups as such can be morally responsible for the effects of their acts, in a sense that cannot be reduced to judgments about individual members’ acts.” (Pettersson, 2008, p. 244). I contest this, arguing that in such cases we

⁴ One notable instance of using game theory to shed light on moral questions, which influenced my thinking, is Runciman and Sen (1965).

⁵ I take this line, aware that philosophers have tried to construct notions of collective intentionality (see Searle, 1990), which has to be taken seriously. However, analyzing these ideas using a game-theoretic lens, they do not stand up to scrutiny (for a philosophical critique, see Zaibert, 2003). We can of course talk of shared preferences, as exemplified by expressions like “we prefer” (Sugden, 2000), but that is distinct from group agency, for which, at least for now, I do not see a meaningful definition.

⁶ For analyses of moral responsibility of group entities, such as corporations and universities, see Copp (2006), Pettit (2007), List and Pettit (2011).

must not hold the group morally responsible, even though we may create penalties and rewards to alter behavior. The need is to draw a distinction between different kinds of responsibility, of which *moral* responsibility is one (Tannsjo, 2007).

2. Two Tales

2.1 After the Seminary Game

Consider the interaction between two individuals, Johnny Won and Jaya Tu, whose names, luckily, can be shortened to 1 and 2. They will be called ‘players’, since these are people who choose between alternative actions. In this ‘game’, player 1 chooses between actions that we call A and B, and 2 chooses between actions that we call A and B. Hence, the ‘outcomes’ of the game are (A, A), (A, B), (B, A), (B, B). The earnings of the players are shown in the payoff matrix labelled “Before the Seminary” in Table 1, where 1 is the row player and 2 is the column player. Player 1 chooses between the rows and player 2 chooses between the columns. The payoffs in each cell are, respectively, player 1’s payoff and player 2’s payoff

If 1 chooses A, and 2 chooses B, both of them earn 100. If 1 and 2 both choose A, they earn 100 and 101. Finally, in game-theoretic language, if they choose (B, A), they earn (101, 100) and if they choose (B, B), they earn (101, 101). The consequences of their choices are fully described by the payoff matrix Before the Seminary. For simplicity we may think of the payoffs as dollar earnings.

Given each individual’s wish to maximize his or her own earnings, an ‘equilibrium’ or ‘Nash equilibrium’ is an action pair from which neither player has an interest in deviating unilaterally. It is easy to see that the only equilibrium is (B, B), where they earn (101, 101). If 1 deviates to A, he earns 100 and if 2 deviates to A, she earns 100. Hence, neither will wish to deviate and so (B, B) is an equilibrium of Before the Seminary game. Suppose that a triple digit income is a very high income, so note that, no matter what the players do, they are very rich.

Table 1. A Society

Before the Seminary

	A	B
A	100, 101	100, 100
B	101, 100	101, 101

Bystander’s Earning’s Matrix

	A	B
A	2	8
B	0	4

Now suppose, Johnny Won, for reasons we need not go into here, goes off to the seminary, where he learns the importance of morality. A minimal requirement of morality is to be mindful of how one’s actions affect the poor. When asked if he gave thought to this, he admitted that not only did he not, in fact he had no idea how the poor bystanders were affected.

As it happens, this society has one bystander. He is very poor and entirely dependent on what 1 and 2 do. In particular, if 1 and 2 choose actions (A, A) the bystander gets 2. If they choose (A, B), she gets 8. If they choose (B, B) or (B, A), she gets 4 or 0, respectively. All this is summed up in the Bystander's Earnings Matrix in Table 1. Henceforth, the collection of the two players and the bystander and their choices and payoffs will be referred to as a society.⁷ Table 1 fully describes this society.

After Johnny Won returns from the seminary, he is a moral person, and he gives equal weight to the poor bystander's earnings and his own. As a result, the game converts to the 'After the Seminary Game' shown in Table 2. This is created by adding the bystander's earnings to 1's earnings, since that is what player 1 now tries to maximize.

What will the outcome be in this game? It is easy to see, it has only one equilibrium, (A, A). The old equilibrium (B,B), is no longer an equilibrium since if player 2 chooses B, player 1 will choose A in order to help the bystander. That is in his best interest as a moral person. But knowing that player 1 would choose A, player 2 would choose A. What happens to the bystander? Note that before player 1 went to the seminary, the bystander had earned 4. After 1 returns from the seminary, the bystander's income is reduced to 2. The concern that player 1 learns to have for the bystander changes the incentives of both players 1 and 2 in a way that leaves the bystander worse off than before Johnny became moral.

Table 2. After the Seminary

	A	B
A	102, 101	108, 100
B	101, 100	105, 101

Outside observers, seeing the atrocious behavior of this group of two rich persons pushing the poor to a low income of 2, will find it difficult to understand that this is happening after one of them becomes moral. Clearly, morality on the part of an individual, *in a strategic environment*, does not necessarily translate into a moral outcome.

The observant reader may notice that if all players became moral (that is, each player seeks to maximize the sum of what he or she earns plus what the bystander earns) the game would end up at the morally best outcome, (A, B), with the bystander earning 8. This may give the impression

⁷ In strict, game-theoretic terms we could think of all of them as players, even though the bystander has no agency.

that if all members of a group become moral, then the outcome is always moral. Alas, that is not the case.

2.2 The Samaritan's Curse

I present now a game with two players, 1 and 2, as before, but each with 3 actions, A, B and C, to choose from. Hence, the society has 9 possible outcomes: (A, A), (A, B), (A, C), (B, A), and so on. The players' earnings are shown in the payoff matrix, labelled The Basic Game in Table 3. Assuming that the players are selfish, it is easy to see that this society will end up at (B, B). This is the outcome from which no player has an interest in deviating unilaterally, and so it is the equilibrium.

Table 3. Another Society

The Basic Game

	A	B	C
A	102, 102	80, 120	108, 108
B	120, 80	104, 104	80, 102
C	108, 108	102, 80	106, 106

Bystander's Earnings Matrix

	A	B	C
A	20	4	0
B	4	6	10
C	0	10	4

Note that in this interaction neither player paid any attention to the fallout of their behavior on bystanders. Bystanders are marginalized people, whose well-being depends on the choices made by the players. Assume, as before, that there is only one bystander, and the payoff to this person is shown in the 'Bystander's Earnings Matrix' in Table 3. Since the equilibrium is (B, B), the bystander gets a payoff of 6 in equilibrium.

Now suppose a good Samaritan comes to town, is dismayed by the moral degeneracy of the people, and gets down to teaching players basic morals. Morals can be of many kinds. There is a growing literature in economics on this (Bowles, 2008; Sen, 2018; Alger and Weibull, 2019; Blume, 2019; Roemer, 2019). The Samaritan focuses on a consequentialist ethic--a kind of utilitarianism⁸ with attention to equity. He tells them: Your choice of (B, B) gives you (104, 104), but do you not see this leaves the bystander with a miserable payoff of 6? If you, the super-rich, opted for (A, A), you would each lose only 2 and the bystander would get 20. Surely you should be prepared to sacrifice \$2 for an additional \$14 for the poor bystander? Ignore the other player, if you wish, since she is super-rich like you, but be mindful of what happens to the poor when you choose.

Suppose both players become moral creatures who value the earnings of the bystander (or of anyone who has an income of less than, say, 25) as much as his or her own payoff. If the outcome

⁸ For a description of consequentialist and non-consequentialist public persona, see Gintis (2015).

is, for instance, (A, B), player 1 gets a consolidated payoff of 84, which consists of 80 for herself and 4 for the poor bystander. By writing in the consolidated payoffs for both players, we get a new game, the ‘The Samaritan’s Curse,’ shown in Table 4. The equilibrium of this new game is (C, C).

The players’ turning into good Samaritans worsens the bystander’s condition. The bystander was earlier getting 6, and now gets only 4. What we have is a kind of Prisoner’s Dilemma or Traveler’s Dilemma. In reverse. In those games, rational behavior leads to the worst outcome. In the Samaritan’s curse, transforming players into good Samaritans leads to a new Nash equilibrium that hurts the poor bystander. In this *individually* moral behavior makes the *group* behave immorally.

Table 4. The Samaritan’s Curse

	A	B	C
A	122, 122	84, 124	108, 108
B	124, 84	110, 110	90, 112
C	108, 108	112, 90	110, 110

So it is now clear that neither one person turning moral nor the whole group turning moral assures an escape from this paradoxical result⁹. These two games warn us that when we look at a group’s immoral behavior, and implicitly think of the individuals constituting the group as immoral, we may be completely wrong.

Some may object to the above analysis by arguing that in both games, After the Seminary and the Samaritan’s Curse, the players described as moral are not truly moral. If they were, they should analyze the consequences of their behavior and see that their morality is doing harm. This should prompt them to change their behavior—be selfish in order to achieve a generous outcome. The problem with this is that one would always be tempted to cheat and play unselfishly after having duped the other person into thinking that one is selfish. But, knowing this, the other person will not be duped. Consider the game After the Seminary. Player 1, after becoming moral, has good reason to be selfish and choose B. This will make player 2 choose B, and not hurt the bystander as much as moral behavior on the part of 1 does. However, if the

⁹ One way out, if everybody has the morality, is to choose from behind a veil of ignorance, so that you do not know who you will be in the game. Such a game would be described as a unanimity game (Basu, 2010). In that case the problem gets resolved in the sense that the outcome that gives the highest payoff is always part of an equilibrium. Bill Gates once wrote (*Economist* magazine, 16 February, 2019, p. 62, column 1) about the importance of going beyond self-interest, on how we ought to have moral causes that we stand up for: “If you want to improve the world, you need something to be mad about.” What we just showed is that if everyone takes this to heart and is mad about a larger cause, the world may end up in a worse place. The only way out is for everyone to be mad in the same way.

outcome of the game is headed towards (B, B), player 1 should cheat and make the moral move, that is, choose A, because that way the game will end up at (A, B). The trouble is player 2 will surely know this and so choose A; and we are back to (A, A). This same problem arises in standard game theory with selfish players. When a certain behavior does harm to all individuals, why don't individuals change their own behavior? This has been discussed in the context of individual rationality and there is no easy resolution of the problem by this route (see Basu, 1994).

What the above games do is to remind us that, when it comes to policymaking, we need to do for moral behavior what we have done for rationality. A major contribution of game theory was to demonstrate that individually rational behavior may not lead to rational outcomes for the group. This has led to thinking about how we as a collectivity can introduce laws, taxes and rewards to bring individual behavior into alignment with our group interest. This kind of thinking has led to a host of tax policies for preserving the commons and global agreements on climate change. The two games above prompt us to think in similar terms to bring individual morality into alignment with group morality.

It is easy now to see how popular discourse on moral responsibility is often flawed. During British rule in India, there were many Britishers who said that their aim was to help develop India. This is commonly treated as hypocrisy because there is evidence that the Indian economy was drained during colonial rule. In other words, behavior by the British players left the bystanders, the Indians, exploited. But putting it in those words immediately alerts us that, by the logic of the above games, we cannot infer from the outcome of exploitation an intention by all the British rulers to exploit India. Before moving on, I hasten to add that for something to be possible does not mean that that is the case.

The above analysis sheds light on vexing issues of group moral responsibility. When a society behaves badly, should we hold the individuals in that society responsible? The answer is no. In the Samaritan's Curse surely we will not hold the two players morally responsible for the harm to the bystander, since each player takes the bystander's earnings into account equally with his own. Do we need to distinguish between group moral responsibility and 'plain responsibility'? Is there a general definition of those who are exempt in both senses? The questions are answered below.

3. The Pulwama Conundrum, Moral Responsibility and Tort Liability

To begin with the last question, let me recount the Pulwama incident mentioned in the introduction. In early 2019, a convoy carrying security personnel on the Jammu-Srinagar Highway, in India's Pulwama district, was blown up in an unprovoked, peacetime attack by a suicide bomber, killing 46 members of India's Central Reserve Police. The terror-group operating out of Pakistan, Jaish-e-Mohammed, took responsibility for the attack. The 'Pulwama attack' was condemned around the world, including by Pakistan. Speaking on Indian television, a day after the attack, the Indian politician and former cricket celebrity, Navjot Singh Sidhu, condemned the attack as "cowardly and dastardly." Then he went on to ask, in a spirit of caution: "Can an entire nation be blamed for a handful of people?" Sidhu's hesitancy in spreading the blame caused

furor. Sony TV asked Sidhu to be taken off the television show. The incident raises questions about who takes moral responsibility when a dastardly human-made event occurs. This real-life controversy raises questions similar to those raised in books and journals of philosophy.

The simplistic rule that the responsibility of some individuals committing a crime can be transferred to the group to which these individuals belong is, I would argue, fallacious. To see this, think of a larger group than Pakistan. Consider, for instance, the group consisting of all Pakistanis and all the Brahmins of India. Call this new entity Brakistan, the individual members being Brakistanis. By the simplistic rule described above, the moral responsibility of the dastardly attack lies with all of Brakistan. Surely, no thinking person will take such a line.

One possible objection to my argument, drawing on philosophical debates on mereology, is that a group is not an arbitrary collection of parts. There must be some common trait. But that is easily amended. Some people claim that Pakistan is behind much of what happens in India. And other people claim that the Brahmins are powerful and behind much of what happens in India. Hence, Brakistan is a group such that for each of its parts there exist people who believe that the part is behind much of what happens in India.

There are people in Pakistan, such as remote tribals, who have no idea they are Pakistanis. There are Pakistanis who consider such attacks abominable, and would like to be able to take steps to stop them. Surely, the moral responsibility for the immoral act cannot be placed on these people, just as it is absurd to hold the Brahmins of India morally responsible for the attack. In brief, even if we believe that moral responsibility for an attack by some can be transferred to a wider group, the group has to be carefully pared down to a meaningful collection.

How is this task of paring to be formally performed? I propose a minimal taxonomy of moral responsibility which enables us to do this. Begin with the set of all the people in a society (or even on earth). For each such person, there will be a set of actions from which this person can choose. After each person makes a choice, we have a vector of choices, one for each person in society. We refer to this vector as an 'outcome'. Let S be the collection of all possible outcomes.

Define a 'bad eventuality', B , to be a nonempty subset of S . It could be the collection of outcomes which entail war, or bystanders getting an income below subsistence, or a terrorist attack. Since I am using a binary characterization, it is harmless to refer to all outcomes that are not bad as 'good'. Our problem is to apportion responsibility for the occurrence of a bad eventuality.

In discussing group responsibility, there has been a lot of discussion on what constitutes a group. To be held responsible, do the members of the group have to have consciously joined the group or more minimally be conscious of being part of the group? (Held, 1970; Tannsjo, 2005). Depending on the specific problem we are discussing, it seems possible to give different answers. I would argue that the essential requirement for 'moral responsibility' for a bad eventuality is whether or not the agent in question could have stopped it. Carrying this idea to the above formal structure gives us a definition of both moral responsibility and the absence of any responsibility.

I define a person i as having ‘*no responsibility*’ for a bad eventuality if, for every vector of individual choices that is a bad outcome (that is, it belongs to the set, bad eventuality), no unilateral change in person i ’s choice can change the bad outcome to good¹⁰. In holding a group as ‘morally responsible’ for a bad eventuality, we must, minimally, leave out all individuals who have *no responsibility*.

In the Pulwama problem, this would entail leaving out individuals who could have taken no conceivable action *unilaterally* to stop it. This is not to suggest that describing the exact contours of this set will be easy¹¹. But this argument shows that it is clearly wrong to blame any group, such as Brakistan, simply because that group contains the individuals who planned the massacre.

Turning to the other extreme, it is easy to define a set of people who have strong moral responsibility for a bad outcome. I define a person i as having ‘*strong moral responsibility*’ for a bad eventuality, B , if there exists an action available to i such that the good eventuality is assured if i chooses that action, no matter what others do. In this case i can single-handedly stop the bad eventuality, *without even having to conjecture what others may choose*.¹²

Given a bad eventuality, denote the set of people with no responsibility by R and those with strong moral responsibility by M . If N is the set of all people in society, clearly, R and M are subsets of N with no intersection between R and M . Further, there may be people who belong to neither R nor M , thereby creating scope for defining intermediate concepts of responsibility. There is an interesting parallel here with some work on cognitive science and artificial intelligence (see Halpern and Kleiman-Weiner, 2018). As Chocker and Halpern (2004) observe, when we think of an individual causing an outcome this is usually an all-or-nothing concept. But as soon as multiple agents are involved, the possibility arises of responsibilities of different degrees. To take their motivating example, if B wins an election against G by a vote of 11-0, each of the voters have much less responsibility for the outcome than when B wins by a vote of 6-5.

I want to put in a word of caution for scholars who venture to create intermediate notions of moral responsibility. Suppose in a game with 3 or more players, no player can unilaterally prevent a bad eventuality (where there is a terrorist attack, say) but there is scope for 2 players to take a coordinated action (they punch the attacker simultaneously) which would prevent the bad eventuality. Some may wish to hold these two players morally responsible for the bad outcome. I am not persuaded by this. The fact that coordinated action by the two can stop the bad eventuality does not seem to be a matter of any significance in itself. To be able to coordinate

¹⁰ We can think of the bystanders as players, with one action to choose from, and it is immediate that they have no responsibility for a bad eventuality.

¹¹ Underlying this is a deep matter concerning the mapping of real-life problems into games. As Samuelson (2016) noted, the importance of game theory in real-life occurred as classical game theory gave way to an “instrumental view”. This entails translating real-life situations into games and then analyzing them using the tools of game theory. What is important to recognize is that the translation of complex reality into a game does not occur by any formal rule, but by intuition and commonsense, and there will always be scope for contestation here.

¹² To draw on an example from Chocker and Halpern (2004), if a candidate wins an election by a 6-5 vote, we can say that each voter has full moral responsibility for the candidate’s victory.

you need to be able to speak or communicate. Someone, for instance, has to say, ‘Let us both punch the attacker at the same time.’ If in reality it is possible to make such statements that should be part of the game. If it is not, then the fact that the two players *could have* coordinated seems to be irrelevant.

There, nevertheless, are ways to go further in apportioning weaker forms of moral responsibility. In the paragraph above where I define strong moral responsibility, a part of a sentence is in italics. That gives a hint of possible intermediate definitions. One option is to define a player as having a ‘weak moral responsibility’ for averting a bad eventuality B, if for every outcome that belongs to B, there is an action on the part of person *i* such that if *i* unilaterally deviated to that action, the new outcome would be outside of B.

One reason to be cautious about this definition is that a player has to be able to predict what others would do to enable society to reach a good outcome. There are however contexts where this ability to predict is plausible. Consider a society with only one equilibrium, which happens to be bad. Now suppose there is a player who by the above definition does not have strong moral responsibility for the bad eventuality. But suppose there is an action such that if, when the equilibrium is to occur, he or she deviates unilaterally to this action, then the outcome would be good. We may wish to hold this person morally responsible in some sense. This is because the outcome that would happen in such a society (with a unique equilibrium) is not hard to see.

To see the terms defined above at work, consider the After the Seminary Game. Let us suppose that any outcome in which anybody gets a payoff less than 3 is part of a bad outcome and a bad eventuality is the set of all bad outcomes. Hence, the outcomes (A, A) and (B, A) are the bad ones. In this case, if the bad eventuality occurs, player 1 has no responsibility, for it occurs whether he does action A or B. But player 2 has strong moral responsibility: if she does B, no bad outcome occurs. In other words, when after the visit to the seminary, player 1 returns to the society and plays the game and the moral outcome deteriorates, the moral responsibility for this bad outcome lies entirely on the shoulders of player 2. The collective immorality is not the responsibility of the person who went to the seminary but of player 2.

In the Samaritan’s Curse, if we suppose that only the outcome (A, A) is good, and all the rest are bad, then clearly neither player has strong moral responsibility for the bad outcome—it requires action by two people to avoid it. At the same time, no player is fully absolved of moral responsibility. Only the bystander in this society has no responsibility.

Similar issues arise with Feinberg’s (1968) famous example of a person drowning on a beach with 1000 sun-bathers, each of whom could have saved the man. Every sun-bather has strong morally responsibility, as we have defined it¹³.

¹³ There is an interesting inversion problem to this. Suppose a man who is about to drown is saved because all thousand sun bathers rush to help. It is Suzy who jumps in moments before others and saves him. Note that even if Suzy had not done anything, the man would nevertheless have been saved because the next person would save him. Should we then say that Suzy did not save him since there is no causal link between Suzy’s trying to save him and his being saved? This is an analogue of a problem discussed by Hall (2004).

When it comes to punishment, however, other issues crop up. As Feinberg points out, under the common law none of the sun-bathers is liable for punishment. The logic is that, if you held them liable for punishment, you would have to punish all 1000 sun-bathers and that does seem absurd. There is however an interesting problem of numbers that crops up here which Feinberg overlooks. Let n be the number of sun-bathers who did not do anything. If n is sufficiently small, say 1, 2 or 3, it seems reasonable to punish all, but if n is large, one thousand or one million, most of us would go along with the common law prescription. It is interesting to note that people act on this intuition. In the bystander effect, the greater the number of people who witness someone in need of emergency help, the less likely an observer will take action.

It is important to note that punishment and moral responsibility are not two sides of the same coin. In Feinberg's example, no matter how large or how small n is, all the sun-bathers have strong moral responsibility by the above definition, even though they are not punished under common law; and many of us would agree that if n is sufficiently large, they should not be punished.

Another class of problems arises, akin to what happens in the Samaritan's Curse, if the only way to save the drowning man is for all (or a sufficiently large number of) sun-bathers to take some special action 'Save'. This is similar to the standard pollution control problem or protest marches. Suppose in a big city pollution will drop if and only if everybody takes individual action to control pollution. Further, individual action being costly, it is worthwhile for each person to take such action if and only if such action results in pollution dropping. Sen (1967) called this the Assurance Game. A similar problem arises with rebellions, such as the anti-authoritarian movement in Hong Kong, as in the paper by Cantoni, Yang, Yuchtman, and Zhang (2019)¹⁴. For many of these problems, there will be two equilibria—one in which a large number join in and another in which no one does.

Suppose we are currently in an outcome where no one is taking action to control pollution¹⁵. Whose responsibility is the pollution? Is it right to punish collectives that have not deliberately done something wrong? (Tanssjo, 2005). In this example no one is *morally* responsible. No single individual can change the pollution that prevails. On the other hand, no one is absolved of all responsibility. This is where we may decide to have punitive taxes. As far as punitive actions go, there are situations in life where we impose taxes not because you have done something wrong but because the tax can alter the game and hence the outcome (Basu, 2018). Its justification lies in "compensation and prevention" in the spirit of tort law, instead of "punishment and dessert." In the above example, if we impose a tax for not taking action to control pollution, this could change the game in a way that taking action to control pollution becomes the unique equilibrium.

¹⁴ Because in reality, it is not clear what the game is since the payoffs are not public, in the case of the Hong Kong protest there is an open question about whether you get innate pleasure from joining in the protest, when others do. By doing a carefully controlled study they establish that the game of rebellion exhibits strategic substitution.

¹⁵ In the event of multiple equilibria, which equilibrium will occur may depend on social and psychological cues. While this is an exciting subject for understanding what happens in reality (Basu, 2018), I stay away from that debate here.

The above definitions allow us to formalize Petersson's (2008) distinction between *liability* and *collective moral responsibility*.

In brief, there are contexts where we may need to punish individuals without having to hold them morally responsible. A lot of law relies on this. Peer monitoring in credit markets, whereby if one person defaults, others in her group are punished, as happens in Bangladesh's Grameen Bank, is an example of this (Stiglitz, 1981). It is a punishment justified in a purely teleological way—it achieves a desired end result. By way of digression, I may point out that peer monitoring has a long history. It was used quite effectively by James Mill. He followed the rule of punishing his son John Stuart, whenever John's siblings behaved badly. The recognition that a group can achieve immoral outcomes with no individual acting immorally opens up avenues of analysis which can help us design interventions in the same way that the recognition that a group of rational individuals can destroy their own environment has made possible improved environmental policies.

In the present context, consider North Korea. We do not like what the North Korean leaders are doing. However, we typically we do not allow for the thought that no North Korean leader may want to do what the North Korean leaders are doing collectively. We should consider this. It is possible that Kim Jong Un may hate the outcome of the leaders' collective behavior, in the same way that the players of the Samaritan's Curse game must feel, but he has no way out. He may fear for his life if he were to unilaterally try to change course for North Korea. He may thus be in a trap the same way that Havel (1986) conjectured that the leaders of a post-totalitarian state may have no unilateral exit, or that Akerlof (1976) hypothesized that the upper caste oppressors in a caste-based society might be in a trap.

Once we open this avenue of research, we may be able to analyze instances of collective immorality very differently and devise new kinds of interventions. In the case of nations like North Korea, the US may develop an island in the Bahamas region, with luxury homes and high security, and make offers to leaders like Kim Jong Un that, if they relinquish power, the US will evacuate them out of Pyongyang and settle them on the 'Dictators' Island'.

4. The Game of Life, the Moral Intention and Some Open Questions

This section comments on two open-ended matters in the hope of prompting future inquiry. I begin with an area of inquiry where the problem highlighted in this paper crops up but has not received enough attention: corruption in large organizations. Just as it can be difficult to pin down how to distribute the responsibility of an immoral group act on the individuals that constitute the group, for certain kinds of corruption involving large corporations it can be difficult to decide who is responsible. Much attention has been paid to principal-agent characterization of government. In this relation, the focus has been on the principals, that is, top echelons of government that have the authority to control corruption and the agents, namely, the government officials, who interact with ordinary citizens and can indulge in corrupt practices (Shleifer and Vishny, 1993; Basu, 2018). However, with the rise of the big corporation, there is a similar structural complexity on the side of the bribe-giver or the rule-breaker. Investors often grant managers contingent long-

term incentive contracts to ensure that managers serve the investors' interest. Similarly, managers create an environment for junior personnel so that the junior staff do what serves the managers' interest.

In these complex structures, when there is bad corporate behavior, such as acts of corruption or criminality, it may not be easy to apportion the blame, for reasons similar to those illustrated in the games above. Investors and senior managers can deliberately create structures where the criminal or corrupt deeds are, on the face of it, committed by the junior staff, so that the senior staff and investors are largely absolved. But in reality the corporate structure may be designed in a way that virtually compels the lower staff to commit the act of corruption. In the language of the above games, we have to be alert to the possibility that there may be someone who designs the Samaritan's Curse game to be played by the two players because he or she has a vested interest in the outcome (C, C). Some of this is probably already happening in the corporate world. The problem will get more serious as robots and autonomous AI agents begin to take on more tasks, including decision-making (Halpern and Kleiman-Weiner, 2018; Friedenbergr and Halpern, 2019). There will then be even more complex problems in assigning guilt for bad corporate behavior and, this in turn, can be taken advantage of by senior management and investors to blur the lines of the responsibility. We will need conceptual clarity and intelligent regulation to prevent or at least discourage such practices. The games described above alert us to the fact that responsibility for corruption in a large organization can be complex and is a matter that needs future research.

I have thus far skirted around the dictum, attributed to Kant: 'ought implies can'. If an agent cannot do something, the agent cannot have moral responsibility for it¹⁶. In popular discourse we often use the ought injunction on groups: America ought to propagate peace, the Turkish government ought to allow free speech, the North Korean leaders ought to make way for an elected group. In each of these cases, the Kantian dictum implies that the group in question can do what they are being urged to do. But how do we define what a *group* can do?

In the game of life, each player has a feasible set of actions or strategies, which is meant to be the set of all the 'cans' for that player¹⁷. Without questioning where that comes from, let us ask what is the set of all the 'cans' for the *group* of all the players? A natural answer is the Cartesian product of all the sets of actions available to all the players. There is reason, however, to contest this. Since a group has no agency, what a group 'can' do has no natural definition. If a particular

¹⁶ This is not an uncontested claim. Tessman (2015), for instance, argues that if you do not save a drowning child because there is no way for you to do so, you nevertheless have done a moral wrong. Tessman is a skillful analyst and provides powerful arguments to defend her line. My view is that if we accept her argument, there will soon arise the need for having two kinds of oughts: ought 1, which implies can, and ought 2, which does not. It is also arguable that Hume did not subscribe to the dictum 'ought implies can'. This is for reasons of doxastic involuntarism, namely, the axiom that there is an essential involuntariness behind our beliefs (Qu, 2017).

¹⁷ There is an open question here about determinism. If we human beings are determined creatures, by our genetic inheritance and environmental factors, then what does it mean for us to have choice? I have argued in Basu (2000) that you can be both fully determined and be able to choose. It is just that what you choose is predictable by your genes and environment. So there is no conflict between using the Kantian edict and being fully determined.

outcome occurs, one reasonable assertion is that all the outcomes that could occur through unilateral deviations of each player are the set of 'cans' for this group. This is a methodological individualistic definition of the feasible set, based on what actually happens. Another option is to assert that the set of all the equilibria of the game is the set of all that can happen. But to assert that an outcome which is not an equilibrium and which entails coordination among all the players is something the group can do is deeply problematic.

In practice, when we talk of group agency we usually do something which is logically flawed, but may well be the only way to proceed. Consider the game of life, that is, the ultimate game that is possible, where all individuals are accounted for and we allow all persons to choose freely with no restrictions. Strictly, there should be no scope for discussing further how a bad outcome can be avoided. This is because if such discussion were possible, then it should already have been a part of this game. Yet we do this all the time: describe the game of life and then step beyond it and discuss how we can avoid some bad outcome.

The reason why we can do this is that there is, strictly, no such thing as a game of life. This is so in the same way that there is no such thing as a set of everything. The non-existence of the game of life in fact follows from the non-existence of the set of everything (Basu, 2018). However, in practice, just as in mathematics the universal set is an artificial construct, an agreement among mathematicians that we shall not step outside of this, the game of life can be an artificial construct, an agreement among analysts to work within its confines¹⁸.

What we often end up doing, and it is not clear there is a way out of this, is to do our formal analysis within the confines of the game of life, and then step beyond it to discuss what we can do to avoid bad outcomes. This brings me back to what was briefly touched on in section 2. Seeing that being moral leads to an immoral outcome in the Samaritan's Curse, can the players not decide to change their behavior?

The discussion of how they can do so lies beyond the game because the game did not include talking and changing behavior as among the actions available to the players. What I am claiming is that this contradictory practice of describing a game of life and then stepping beyond it is important. We should not wait until we learn how to resolve the contradiction. We must continue to step beyond it.

This makes it possible to accommodate an intriguing suggestion put forward by Putnam (2005, p.24), drawing on the work of Emmanuel Levinas: "For Levinas, the irreducible foundation of ethics is *my* immediate recognition, when confronted with a suffering fellow human being, that

¹⁸ The game that we actually play and the sets of actions we perceive as available to us may be shaped by our collective thinking, including by the laws we adopt. While the conventional view of law and punishment is that of interventions that change the payoffs from certain actions, in reality they can change the meanings of certain actions. This can lead to the closing off of certain actions or even making us aware of actions that were earlier unutilized because they were beyond our awareness. This line of inquiry, picking up on the expressionist function of the law (Sunstein, 1996) can lead to novel ways to conceptualize how new laws may impact the very conception of the game we play and, through that, the actions we choose (see Hoff and Walsh, 2019).

I have an obligation to do something. [Even if I cannot actually help,] not to feel the obligation to help the sufferer at all, not to recognize that if I can, I must help ... is not to be ethical at all." Note that Putnam (and, for that matter, Levinas) is not contradicting the dictum 'ought implies can'. He is not asserting you ought to help someone whom you cannot help (because it is beyond what is feasible for you) but that you must feel the obligation to help. This of course takes us beyond the boundaries of the game of life since "feeling the obligation to help" was not a part of the game.

Building on Putnam's suggestion, I would argue that the true lesson of games like After the Seminary and the Samaritan's Curse is to recognize that even if a good outcome is beyond our reach, we must nurture and keep alive what may be called the 'moral intention,' which is the intention to achieve the good eventuality. Thus when society gets caught in the equilibrium like (C, C) in the Samaritan's Curse, the morally good outcome, namely, (A, A), is arguably beyond what can be achieved. The achievable outcomes are (A, C), (B, C), (C, C), (C, A), C, B) and they do not include (A, A). This is because going to (A, A) entails coordinated behavior. Someone has to initiate the coordination but that is not part of the game.

To put it differently, in the Samaritan's Curse with moral players, the outcome will be immoral. There is no escape from this. Hope in such situations must lie in the moral intention, the capacity to recognize that the outcome is morally unacceptable, or, in the words of Putnam, to feel the obligation to do more for the poor bystander, even if one cannot actually do so in any immediate sense.

If the moral *intention* is deadened, and we end up in a morally-bad outcome, we would treat this as a *fait accompli* and do nothing. It is our moral intention that makes us want to step beyond the game of life and think of how we may alter our behavior, such as by adopting deontological ethics (do A when playing a game like the Samaritan's Curse, without thinking of the consequence) or altering the rules of the game by imposing taxes and punishment on players to help break out of the dismal outcome. The need is for deontological ethics for *individuals*, which are derived from consequentialist concerns for the *group*. We have seen this in the context of long-run survival of groups from research that shows how a group is stable in an evolutionary sense when individuals act not purely out of self-interest but a convex combination of self-interest and morality (Alger and Weibull, 2013; 2019).

There should be scope for carrying some of these ideas to other domains. Some of the problem arises from an innate intransitivity in many moral matters, whereby each step takes the group to a higher level of welfare but all the steps together make them worse off, as was pointed out by Parfit (1984). By bringing elements of game theory into this can help us formalize this intuition (Basu, 2000; Voorneveld, 2010), which suggests that we may need to commit ourselves to rules of moral behavior which on each occasion may seem to conflict with the collective aims of society but, followed by all, enhances collective welfare. Analyzing behavior beyond the game of life takes us to uncharted territory. Conventional analysis, within the confines of the game of life, runs out of ammunition here. But that must not be reason to stop.

5. Two Cautionary Notes

History is replete with misconstrued policies resulting from misreading academic writings. To guard against this, I close with two policy alerts.

First, much of the paper was concerned with making us more circumspect in holding individuals morally responsible for collective immorality. However, there is a case for distinguishing between holding people morally responsible and *saying* that you hold them morally responsible. There are contexts where the former is wrong and the latter serves a purpose. As we already saw, there are situations where no one is morally responsible for a bad eventuality but where it is nevertheless right to impose a punishment or tax for certain kinds of behavior because that can change the outcome for the better. Now, we all know that to be criticized for being immoral is not something people like¹⁹. It is a kind of punishment. In other words, if people are described as immoral for making certain choices, this may change their behavior so that the group makes better choices. This means that there may be a consequentialist reason to describe people as immoral for certain kinds of behavior even when we would not hold them morally responsible for an outcome. Further, certain amoral rules developed by society to create order can come to *acquire* moral significance. Thus, there is nothing moral for a society to develop a rule that everybody should drive on the left. However, once this rule comes to be followed, its violation by an individual seeking some momentary advantage is seen as a moral violation. Hence individuals endowed with morality may not violate the law even if that is in their self-interest.

Second, as noted above, we must not presume that all dictators are reveling in what they are doing. Instead, we should consider the policy option of giving some of them a way out, such as the assurance of security after they relinquish power and a home in the Bahamas. However, it must be realized that policies have long-run consequences. Plans like the above may help end Kim Jong Un's dictatorial regime, but it may prompt other individuals with no interest in dictatorship to become dictators, purely as a step to getting that property in the Bahamas.

¹⁹ There is evidence that people are ready to incur costs so as not to walk past a Salvation Army officer soliciting charity for the poor (Andreoni, Rao and Trachtman, 2017). Thus there is a need to distinguish between innate morals and morals that gets triggered by someone else's action—open palm or speech. In the terminology of Ellingsen and Mohlin (2019), there is a distinction between innate passion and 'decency', the latter being taught and depending on the expectations of others. It is known people like to avoid being blamed. While what causes blame can depend on what others do, it is possible to argue that it depends on what is the collective opinion. All this is predicated on the assumption, increasingly gaining traction in economics, that people *are* innately moral. We can destroy this by priming them to respond only to financial incentives appealing to self-interest. But left to themselves human beings have moral anchors, which can in fact create a more efficient society (see Bowles, 2016).

References

- Akerlof, G. (1976), 'The Economics of Caste, Rat Race and other Woeful Tales,' **Quarterly Journal of Economics**, vol. 90.
- Alger, I. and Weibull, J. (2013), 'Homo Moralis—Preference Evolution under Incomplete Information and Assortative Matching,' **Econometrica**, vol. 81.
- Alger, I. and Weibull, J. (2019), 'Morality: Evolutionary Foundations and Policy Implications,' in K. Basu, Rosenblatt, D. and Sepulveda, C. (eds.), **The State of Economics, the State of the World**, MIT Press, Cambridge, MA, forthcoming.
- Andreoni, J., Rao, J. M., and Trachtman, H. (2017), 'Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving,' **Journal of Political Economy**, vol. 125.
- Arruda, C. (2017), 'How I learned to worry about the Spaghetti Western: Collective Responsibility and Collective Agency,' **Analysis**, vol. 77, No. 2.
- Basu, K. (1994), 'The Traveler's Dilemma: Paradoxes of Rationality in Game Theory,' **American Economic Review: Papers and Proceedings**, vol. 71.
- Basu, K. (2000), **Prelude to Political Economy: A Study of the Social and Political Foundations of Economics**, Oxford University Press, Oxford.
- Basu, K. (2010), 'The Moral Basis of Prosperity and Oppression: Altruism, Other-regarding Behavior, and Identity,' **Economics and Philosophy**, vol. 26.
- Basu, K. (2018), **The Republic of Beliefs: A New Approach to Law and Economics**, Princeton University Press, Princeton.
- Bernheim, B. D. and Whinston, M. D. (1986), 'Common Agency,' **Econometrica**, vol. 54.
- Bjornsson, G. (2014), 'Essentially Shared Obligations,' **Midwest Studies in Philosophy**, vol. 38.
- Blume, L. (2019). 'Comments on Alger and Weibull', in K. Basu, Rosenblatt, D. and Sepulveda, C. (eds.), **The State of Economics, the State of the World**, MIT Press, Cambridge, MA, forthcoming.
- Bowles, S. (2008), 'Policies Designed for Self-Interested Citizens May Undermine 'The Moral Sentiments': Evidence from Economic Experiments,' **Science** 320: 1605–1609.
- Bowles, S. (2016), **The Moral Economy: Why Good Incentives are No Substitute for Good Citizen**, Yale University Press, New Haven.
- Cantoni, D., Yang, D. Y., Yuchtman, Y. and Zhang, Y. J. (2019), 'Protests as Strategic Games: Experimental Evidence from Hong Kong's Anti-Authoritative Movement,' **Quarterly Journal of Economics**, vol. 134.
- Chant, S. R. (2015), 'Collective Responsibility in a Hollywood Standoff,' **Thought: A Journal of Philosophy**, vol. 4.
- Chocker, H. and Halpern, J. Y. (2004), 'Responsibility and Blame: A Structural-Model Approach,' **Journal of Artificial Intelligence Research**, vol. 22.
- Copp, D. (2006), 'On the Agency of Certain Collective Entities: An Argument from 'Normative Autonomy', ' **Midwest Studies in Philosophy**, vol. 30.
- Dworkin, R. (1986), **Law's Empire**, Harvard University Press, Cambridge, MA.
- Ellingsen, T. and Mohlin, E. (2019), 'Decency,' Working Paper, Department of Economics, Lund University
- Feinberg, J. (1968), 'Collective Responsibility,' **Journal of Philosophy**, vol. 65.
- Friedenberg, M. and Halpern, J. Y. (2019), 'Blameworthiness in Multi-Agent Settings,'

- Association for the Advancement of Artificial Intelligence:
https://www.cs.cornell.edu/home/halpern/papers/group_blame.pdf
- Gintis, H. (2015), 'A Typology of Human Morality,' mimeo: Santa Fe Institute.
- Hall, N. (2004), 'Two Concepts of Causation,' in J. Collins, N. Hall and L. A. Paul (eds.), **Causation and Counterfactuals**, MIT Press, Cambridge, MA.
- Halpern, J. Y. and Kleiman-Weiner, M. (2018), 'Towards Formal Definitions of Blameworthiness, Intention, and Moral Responsibility,' Association for the Advancement of Artificial Intelligence: <https://www.cs.cornell.edu/home/halpern/papers/moralresp.pdf>
- Havel, V. (1986), 'The Power of the Powerless,' in J. Vladislav (ed.), **Living in Truth**, Faber & Faber.
- Hess, K. (2014), 'Because They Can: The Basis for Moral Obligations for (Certain) Collectives,' **Midwest Studies in Philosophy**, vol. 38.
- Hoff, K. and Walsh, J. (2019), 'The Third Function of Law is to Transform Cultural Categories,' **Policy Research Working Paper**, No. 8954. World Bank.
- Lewis, D. (1969), **Convention**, Harvard University Press, Cambridge, MA.
- List, C. and Pettit, P. (2011), **Group Agency: The Possibility, Design, and the Status of Corporate Agents**, Oxford University Press, Oxford.
- Marschak, J. and Radner, R. (1972), **The Economic Theory of Teams**, Yale University Press, New Haven.
- Parfit, D. (1984), **Reasons and Persons**, Clarendon Press, Oxford.
- Petersson, B. (2008), 'Collective Omissions and Responsibilities,' **Philosophical Papers**, vol. 37, No. 2.
- Pettit, P. (2007), 'Responsibility Incorporated,' **Ethics**, vol. 117.
- Putnam, H. (2005), **Ethics without Ontology**, Harvard University Press, Cambridge, MA.
- Qu, H. (2017), 'Hume's Doxastic Involuntarism,' **Mind**, vol. 126.
- Roemer, J. (2019), **How We Cooperate: A Theory of Kantian Optimization**, Yale University Press, New Haven.
- Rubinstein, A. (1998), **Modeling Bounded Rationality**, MIT Press, Cambridge, MA.
- Runciman, W. and Sen, A. (1965), 'Games, Justice and the General Will,' **Mind**, vol. 74.
- Samuelson, L. (2016), 'Game Theory in Economics and Beyond,' **Journal of Economic Perspectives**, vol. 30.
- Sartorio, C. (2004), 'How to be Responsible for Something without Causing It' **Philosophical Perspectives**, vol. 18.
- Searle, J. (1990), 'Collective Intentions and Actions,' in P. Cohen, J. Morgan and M. E. Pollack (eds.), **Intentions in Communication**, MIT Press, Cambridge, MA.
- Sen, A. (1967), 'Isolation, Assurance and the Social Rate of Discount,' **Quarterly Journal of Economics**, vol. 81.
- Sen, A. (2018), **Collective Choice and Social Welfare**, Expanded edition, Oxford University Press, Oxford.
- Shleifer, A. and Vishny, R. W. (1993), 'Corruption,' **Quarterly Journal of Economics**, vol. 108.
- Stiglitz, J. E. (1981), 'Peer monitoring and Credit Monitoring,' **World Bank Economic Review**, vol. 4.
- Sugden, R. (2000), 'Team Preferences,' **Economics and Philosophy**, vol. 16.
- Sunstein, C. (1996), 'On the Expressive Function of Law,' **University of Pennsylvania Law**

Review, vol. 144.

Tannsjo, T. (2007), 'The Myth of Innocence: On Collective Responsibility and Collective Punishment', **Philosophical Papers**, vol. 36.

Tessman, L. (2015), **Moral Failure: On the Impossible Demands of Morality**, Oxford University Press, Oxford.

Voorneveld, M. (2010), 'The Possibility of Impossible Stairways: Tail Events and Countable Player Sets,' **Games and Economic Behavior**, vol. 68.

Zaibert, L. A. (2003), 'Collective Intentions and Collective Intentionality,' **American Journal of Economics and Sociology**, vol. 62.