Can Twitter enrich the state?

Twitter storms, the rule of law and the individual

Image source: South Park©

Sara Laurijssen

Master thesis Applied Ethics

22 June, 2018

Utrecht University

Supervised by Rob van Gerwen

Second reader: Jan Vorstenbosch

Abstract

This thesis examines the effects of Twitter storms on both the individual and the rule of law. In order to examine the impact of these storms on individuals, relevant examples like the story of Justine Sacco and the #Me-too discussion are reviewed. When individuals become the object of a Twitter storm, a viral discussion with far going consequences for the user(s) of the platform, their rights are violated. Light is shed upon how this can happen online. Next to individual users, this thesis will argue that the rule of law is hurt by Twitter storms as well. The rule of law is described as a set of values that represent the ideal of a state that shelters its' citizens from tyranny and abuse. Twitter storms both benefit and threaten this set of values. Arguments will be made about democracy, fairness and accountability.

Word count: 16.652

Table of contents

1. Introduction

1.1 Research object

In the last decade, Twitter has rapidly grown as a micro blogging platform where everyone can post all sorts of different opinions, ideas, and facts and share their daily activities (Kwak, Lee, Park & Moon, 2010). Users are able to connect with like-minded others and form movements by sending tweets, reacting, re-tweeting and the use of hashtags. These hashtags help to bring others into contact since they make tweets more visible. When people are tweeting about the same topic and use the same hashtag, Twitter gives them the ability to look at all the tweets that were sent containing this hashtag. They can follow people they think of as inspiring, or follow the news. Using social media to engage with people all around the globe is something that has characterised the 2010s.

In a short time, interacting with others online has become an important part of our lives. It is hard to imagine that we would not participate in discussions or find out what is happening around us through social media. However, Twitter and other social media are not only bringing us into contact with people or content that we enjoy. We are also able to communicate with users who are on the opposite side of the debate (Trottier & Fuchs, 2014). This can lead to harsh discussions online, especially when people in the debate begin to form groups. These groups will then isolate themselves from the debate, increasing each other's opinions and moving the discussion to a more private internet page. This isolation can lead to the erosion of empathy and comprehension (Trottier & Fuchs, 2014). Next to such polarisation, individuals on the internet can easily be harmed. For instance, it has never been this easy to bully someone online and spread hateful statements. What has been posted online can hardly ever be removed after content is downloaded by users and servers. The use of Twitter and other social media can have a negative impact on individuals. This became remarkably clear in the case of Justine Sacco.

Justine, a Twitter user and PR executive was travelling to her family in South Africa. Just before she boarded her plane she tweeted: ''Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!''[1] This line would change her (professional) life tremendously over the duration of her flight from New York to Cape Town. What was intended as a joke became the conversation of the day. The tweet was picked up by other users who thought it was racist and retweeted her tweet over 2,000 times. Many people responded to this tweet and wished her ill. They accused her of hate speech, claiming that she was ''a dumb racist bitch'' who ''should be fired and get AIDS'', or worse. By the time Sacco's plane landed in Cape Town her family was furious, no-one wished to be associated with her anymore and she had lost her job. She made numerous apologies, removed the tweet, later her whole Twitter account, but the damage was already done. The rapid internet storm that was caused by

---

[1]   https://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html

people who felt hurt or were upset about this tweet costed her hugely. Her status as a PR executive was damaged. She went from being an unknown USA citizen to a hated celebrity in just 280 characters.

1.2 Problem

It is not the goal of this paper to comment on the content of this tweet and decide if it was indeed an insult. Opinions on the content of the tweet will vary. However, it is questionable if Justine deserved all that happened to her. While she was in the air, she was unable to react to others and explain her hulking line. In most scenarios where only a single hate speech line, intended as a joke, is voiced, no such heavy consequences as Justine suffered, are seen as acceptable. In most countries, suspects are punished with a fee if they are prosecuted. It is worrisome to witness that individuals such as Justine can become the object of people online shaming them as if they were bound to a medieval pillory. This can lead to unwanted outcomes. Individuals were first presented to court and punished by the use of democratically formulated laws. In case like Justine´s they now have to suffer the consequences of Twitter storms instead. The right to defend oneself becomes lost online when an individual is subjected to ''online witch hunt.'' While Sacco was in the air, she was framed by others as an agent who did something wrong. She became the suspect of hate speech and racism, acts that are seen by many as morally wrong and maybe even as a crime. People responded towards Justine, creating a Twitter storm of users who commented and reacted to each other. Motivated by their displeasure, the other users stirred up the discussion, reaching more and more people. When the discussion grew, a snowball effect was created. More and more people read the original tweet or the comments. A hashtag was created (#hasJustinelandedyet) and helped to increase the reach of the tweet. The discussion could no longer be ignored by media and by those who had a relationship with the person that was blamed. They dissociated from her. Something has to be done to clear their own names. Justine, the suspect, has to suffer the consequences of her action. Justice had to be served.

Situations like these form a problem for the rule of law in the sense that they violate some of its central principles like the right to a fair trial or the authority of the law. When individuals such as Justine are being 'punished' for some alleged misbehaviour because different people ask for justice, online consequences can mostly not be undone. Even if these individuals become the suspects of serious investigations and are sent to court, the damage is mostly already done. The social status that this individual had is already lost before an official trial can even start. Consequences that follow from this loss of social status are mostly harsher on the individual than law prescribes. So the actual verdict is too little too late. In addition, when these individuals turn out to be innocent they already have suffered irreparably. Aside from Twitter, these kinds of internet storms could also be happening on Facebook, Instagram or any other platform where people engage and respond to each other. Often the discussion starts on one of the platforms, and then swiftly moves to another. In this thesis the scope lies on Twitter since most cases of internet trials seem to happen via Twitter

(Pfeffer, Zorbach & Carley, 2013). Since this platform gained popularity around 2010, a lot of scientific data has been gathered, which shows how people use Twitter and to what extent. This data will help to clarify how these Twitter storms take place and why. Next to these two advantages of focusing on Twitter, all these social media have similar characteristics. Studying one of them will help to understand what is happening on most platforms.

1.3 Overview of the argument

This thesis investigates if the rule of law is threatened by this phenomenon. In order to do so, it is important to define what the rule of law is and which values this rule safeguards. These values point out what is at stake. This will be done in section 2, where definitions that are used in this paper are explained. Next to the rule of law, the architecture of a Twitter storm is examined. After this theoretical framework, arguments on the effects of Twitter storms are presented. The various negative effects the online judgement of individuals brings about are given, as well as the positive effects. These arguments are divided into three main categories, democracy, fairness and accountability. Thus, this thesis focuses on the main question:

How are Twitter storms influencing the ideal of a rule of law?

This main question will be divided into four research questions that contribute to an answer:

1. What is the rule of law and what values does this rule of law embody? – Section 2

2. What are Twitter storms? - Section 2

3. How are individuals treated online by others and by the platform? – Section 3

4. Does this treatment do justice to the rule of law? – Section 3

One of the roles an ethicist should play in our society is in warning its members of possible harm. In my opinion a philosopher should act as a guard, protecting the values we have formulated together and see as worth pursuing. New technologies and the way that they affect the things we see as precious should be analysed. The internet and online discussion have many advantages. We should, however, not ignore possible negative side effects.

1.4 Author Justification

This thesis focusses on one of the moral questions Twitter storms raise. Different theories of philosophers will be given, mostly theories about law and how the rule of law should be. One of the main authors that will be discussed is Bovens. Mark Bovens is a Dutch professor who has specialised in law philosophy and institutions for open society. As a member of the Dutch Scientific Council for Government Policy (WRR), his work has great influence on Dutch and European law. I chose him as one of the main authors because he wrote different books and articles about the rule of law and how the internet affects it. The focus of these books is exploratory; a summary of effects is given without prioritising one over the other. From this starting point, other theories are taken into account and are analysed. These theories are

not just theories of philosophers. The internet is a widely described phenomenon in literature. By only looking at the philosophical side of the debate rich context about online behaviour, and the meaning of the rule of law and its values may be lost. Theories of social scientists, law theorists and others will be used to describe what is happening online in order to come to a critical analysis.

1.5 Place in the debate

Since the start of the twenty-first century, law philosophers have been analysing the effects of the internet on the rule of law. They found that the internet is challenging our current laws, but also has promising advantages (Banks, 2010; Bovens & Loos, 2002). Current laws are insufficient to respond to the rapid changes society faces in the age of the internet. New questions about property and online property are raised, as well as questions about what sort of behaviour should be accepted online. Where the internet once was the wild west of programmers and hackers, we now see that governments and international organisations are making new laws in an effort to control this public space. Downloading movies for free has become illegal and people can be prosecuted for hacking. Law philosophers expect that our whole system of judgement and rules will be transformed in order to fit our current information society (Bovens, 2003). There will be a revolution of the rule of law, enabling more and more people to contribute to democratic processes. Voting might be done through apps on mobile phones and people can take part in online courts and juries. The gap between the government and its citizens will decrease as soon as internet traffic is secured. The internet will become a common good. It will be accessible for all, provided by the government and (non-) profit organisations. It is the task of the government to make sure that all can participate in these democratic processes (Bovens, 2003).

A revolution of law has not yet taken place. It might be questionable if this should happen and to what extent. Even so, it is important to think about the way we want to equip the internet in order to reinforce democratic processes. Should the internet and platforms such as Twitter be used by the government and can they be used to enhance the rule of law? This thesis will try to shed a light on these particular questions in order to contribute to these theories that were formed at the start of the twenty-first century. Especially with regard to Twitter storms, we must make sure that there is a discussion about how free the internet should be and what behaviour and attitudes of others we accept online. When something like the rule of law might be harmed, it is important to find out what values are being jeopardised and what consequences this harming brings about. It is clear that if individuals lose the right to defend themselves and the right to a fair trial, society needs a response to this phenomenon.

2. Theoretical framework

This section will explain two of the major concepts that are described in this thesis: the rule of law and Twitter storms. This will help to understand what is at stake and in what way the values the rule of law represents are affected by Twitter storms. This section is split into five sub-sections. The first subsection will explain the concept rule of law. After that, the values of the rule of law are given. Then a couple of comments will be made on how the rule of law is affected by the internet. In the fourth sub-section of this theoretical framework, Twitter storms will be analysed. Relevant theories will be presented on how they occur, which people participate and what sort of process they follow. The section will end with a conclusion, summarising the most important elements of this literature study.

2.1 The rule of law

It proves to be hard to find a clear description of what the rule of law is. Philosophers give different accounts on the rule of law and on how the perfect rule of law should be (Jowell & Oliver, 2007; Tamanaha, 2004). Some even say that there is no rule of law since there are many different rules of law, which are all incomplete. Enlightenment philosophers such as Kant, Hobbes and Montesquieu believed that all people had the same rights and that no ruler should be above the law or act in such a way that large amounts of people suffer from hunger or impoverishment (Tamanaha, 2004). Most ideas on what the rule of law is hold elements of their theories. The rule of law should be a system that prevents tyrants from rising to power and gaining full control over the people. The ideals of these enlightenment thinkers became shared by the public and when rights were detained from the masses, democratic revolutions broke out in Europe. Even though most of these revolutions ended in chaos and eventually the installation of new monarchs on their throne, such as in France, values like freedom, equality, the right to educate yourself and fraternity became widely promoted. For example, the democratic revolutions led to the formulation of the first declaration of human rights in France. This document stated that all are equal under the law and should be treated in such a way. Different systems of rules of law were installed in different countries, but they all shared one thing: it became their mission to promote human rights. Citizens were safeguarded from tyranny. These countries, such as Germany or the Netherlands, did not necessarily become democracies however. A monarch ruled and the people were protected from this ruler by a constitution. The idea of a rule of law, or state of law, was established (Tamanaha, 2014).

The states that now had a rule of law differed from each other. The British law philosopher and lawyer Dicey compared these different systems and abstracted values from them (Jowell & Oliver, 2007). From this comparison, values were formed that give a more complete view of the rule of law and what a state with a rule of law should be protecting:

1. Legality, the law must be followed by both citizens and law-enforcement officials. This last group may only act if their actions are in line with the law.

2. Certainty, the law must be known to all and based in substantive fairness. This means that one can only be punished if he or she could have known that an act is prohibited or required.

3. Consistency, all are equal under the law and must be treated alike.

4. Accountability, rules and laws must be made by the public in order to be of the public. All should contribute equally to this process and are therefore responsible for these rules.

5. Efficiency, laws and rules must be made to govern a country in an efficient way with clear rules and protocols (Jowell & Oliver, 2007).

These values are still visible in the rule of law we see today. With the arrival of the internet it has however become hard to speak about the rule of law. Different states with their own adaptations of the rule of law come into contact online. The rule of law has become a shared idea about a system that protects its people from tyranny. No individual or body of government can rise to power or use violence in order to enslave others. To provide this state of law a democratic system is needed, ensuring that all inhabitants of the state can (indirectly) determine its course. This democratic system can only flourish when freedom is provided. Therefore, there has to be freedom of press and freedom of speech. When democratic laws are formed, they must be respected by all. To establish this, state institutions are put into place in order to enforce them. This is done by the police or sometimes the army. This executive power can, however, only use power when laws have been broken. Punishment and forms of violence are controlled by legal acts and regulations. In a rule of law, all have the right to a fair trial when being judged. To ensure that trials are fair it is possible to ask for a lawyer, the procedures and rules for the process are clear, and judges are appointed for life. These are conditions which have to be met in order to provide fair trial to everyone, and make sure that no one is treated differently (Tamanaha, 2004).

2.2 Influence of the internet

Since law philosophers such as Bovens (2003) predicted that the internet would change our rule of law, it is necessary to know what values might be put under pressure when individuals become the victim of an internet storm. Therefore, this section will give an extremely short overview of how the internet emerged and how the internet is influencing our rule of law.

It might feel as if computers and data were always connected, sharing this data and interacting with each other, but this is not the case. Around 1950 the first ideas of computers interacting with one another were used for communication during the cold war (Leiner et al., 2009). Licklider, working for MIT, wanted to create a system in which information could still be sent from point A to point B when the direct connection between those two points was corrupted (for example by a nuclear attack). To do so, multiple connections were needed between devices that could send information. This system was funded by the army, creating the ARPANET, a network that connected multiple computers over the country and later also overseas. This ARPANET would later become the internet as

we know it today. On ARPANET, the first emails were sent, creating the fundaments for person to person communication (Leiner et al., 2009). These new interactions with people we do not see but are made abstract through our screens raised new questions about behaviour and numerous legal issues. It is hard for a user to explain what the internet is, since it has no touchable substance. The internet is an environment that connects different users who can all build and expand the internet. It enables us to interact with others and is therefore a communication medium. No one owns this medium and it is not limited to a state (Leiner et al., 2009).

2.3 Effects on the rule of law

There are different views from which to look at the way the internet is influencing our rule of law. Some argue that it has merely positive effects, others see the internet and online interaction as a threat to our rule of law. There is however some sort of consensus on how the internet affects the rule of law (Bovens, 2003). The coming of the internet brought about four effects on the rule of law:

1. De-territorialising: since the internet is not bound to borders it affects different states and thus different rules of law.

2. Technological turbulence: new innovations follow each other rapidly. It is impossible to oversee, let alone predict and hence regulate, the consequences of new innovations or the social response to these changes.

3. Horizontalisation: markets, networks and their algorithms form new driving forces in the field due to the internet, changing the power that big organisations and governments have.

4. Dematerialisation, information can travel anywhere in split-seconds since it is no longer bound to material. The utter, communication becomes information.

These four effects are especially visible with regard to democracy. According to Pitkin's theory, democracy is a process of representation and reflection (Pitkin, 2004). She argues that politics must always be in connection with the people they represent. Next, the government must be a representation of the different values and norms that live in society. Only then can a state of law function, as only then there is functioning accountability. The internet is both promising and harmful for this process of representation and reflection. The internet can, for example, be a blessing since it is easily accessible to all. It holds a promise of high-quality communication with those who represent the people. One can chat or even WhatsApp with the local government, activists and politicians. This may help to increase the quality of the representation as well as giving individuals the chance to speak up and share their ideas with politicians. It is thus indirectly increasing the degree of reflection. It has been impossible to apply direct democracy for a long time: there were simply too many people. This is however changing. In the near future, the public will be part of political decisions by use of tele-voting once the technology has become safe and accessible to all (Norris, 2001). The American government is already experimenting with these technologies. They use it to

let people who are oversees, for example soldiers out on a mission, vote in elections (Crawford, Mack, Eugene & Gilbert, 2015). Techniques such as these are not widely used yet however, since it is hard to ensure that this voting happens safely.

Even though the internet can enrich our democracy, it is questionable if this is as eligible as it might sound. Sceptics argue that even though more people can and will participate in the political debate, it is debatable if we should let them take political decisions. Lack of expertise and know-how, media influences and manipulation can lead to ill-considered outcomes (Bovens, 2003). There is simply too much information to become an expert in all fields or google all the arguments. Politics might become for those with more time on their hands and a wider field of expertise when the internet determines who can participate and who can't. According to Bovens (2003) and studies about the demographics of Twitter, people with a higher education are represented the most in online discussions. More will be explained about this issue later. Not only are these intellectuals more likely interested in politics, they are also better equipped to write their statements and influence politicians. Switching to a more direct form of political decision making can lead to the exclusion of those who are not able to participate in online discussions that go beyond common knowledge. This would harm the value of consistency. Bovens (2003) predicts that the effects of the internet on the democracy might not be as promising or alarming as the optimists and the sceptics present them. Few people will interact with politicians online. The internet will become an addition to the means we already have to engage in public debate. It will help the media to spread content that can form political opinions or support politicians during elections. This effect is not new however; the internet is merely a new source of it.

So far, the conclusion is that the rule of law is a system of governance in which the law has the highest authority. States that recognise this idea of governing, have laws which prevent tyrants from gaining power and ensure that their citizens can partake in the governance of the state. In order to enable this, they provide democracy, fair trials, and the means needed to participate in both of these institutions. The rule of law is thus more than a system of governance: it is a set of values. These values and individuals living under the rule of law are influenced by the internet both positively and negatively.

2.4 Twitter storms
In the introduction a Twitter storm was described as an event in which a tweet goes viral because of massive response by others. Reactions build upon one another, creating a snowball effect and reaching more and more users and eventually the media. This process can violate individual users; consequences can be unknown or out of proportion. There is however more to Twitter storms. Millions of tweets are posted every single day. Nonetheless, not all of these tweets have the potential to start a storm. Twitter claims to be a platform were all individuals are equal; they can all share their thoughts and vent their feelings. No matter their background, ethnicity or religion, all users will be heard. So, social status does not play a part in the reach of a single tweet (Morello, 2015). What is of importance however, is the number of followers a user has: the more followers one has, the

more people will see one's tweet and the more people might respond to it. Followers mostly consist of people who are known to the users in real life, possible work or network connections, and people that see them as inspiring (Kwak et al., 2010). So if anyone can start a Twitter storm, it is necessary to answer the next question: what is needed to start a Twitter storm?

The first ingredient is emotion. A study in 2012 by Berger and Milkman shows that in order to gain the attention of others, tweets should contain strong emotions. If these emotions are picked up and shared or contradicted by others, there will almost always be a response to the tweet. If these responses are emotional as well, the reach of the message will grow. Not all emotions work equally well. Awe, amusement, anxiety and anger work the best, while sorrow and sadness are not likely to be picked up by others (Berger & Milkman, 2012). The best way to be retweeted is to sweet talk the ones you are directly in contact with or agree with. If you emphasise their feelings or opinions, you will likely be replied to. If you however disagree with others an equally strong emotional tweet is needed in order to be noticed. Extremely negative tweets draw a lot of attention from other users. Hashtags help to expand your reach as well. A catchy, short and easily understood hashtag can move the discussion to a wider public (Morello, 2015). When others are for or against the case that the hashtag represents they will start to use it as well to support their side of the debate. This so-called hashtag activism is a recurring phenomenon. Twitter storms have one or more hashtags. Hashtags draw the attention of the media, mostly when they become a trending topic. Trending topics are topics that dominate the debate. A list of them is displayed for all users that Twitter thinks they are relevant for. They can be found at the home page or next to someone's timeline. When topics are used on an extremely large scale, they can even become visible for all users. This depends on how many people are reading them and adding to them. Events such as Charlie Hebdo or the Brussels bombings became trending topics worldwide. When Twitter storms become massive they are likely to become relevant trending topics (Morello, 2015). This happened for instance with #ShirtStorm and #Shirtgate in the events of November 2014. When a scientist from the Rosetta mission was explaining the history of the comet that they would later try to land a device on, all eyes turned to his shirt, which displayed images of barely clothed women. Twitter exploded and numerous accounts wrote about this shirt, calling it sexist. Questions about women in science were being raised, and the wearer of the shirt, Matt Taylor lost his job at London University. The media had a clear part in this outcome since several newspapers posted about this incident (Morello, 2015).

Aside from these two ingredients to start a storm, Twitter storms are also characterised by the people who engage in them. A study that investigated participants in Twitter storms was done by LeFebvre & Armstrong (2016). This study examines the participants in the Twitter storm after the Ferguson shootings where a black teenager was killed by a police officer. They distinguished different types of participants in the Twitter storm that followed the events (LeFebvre & Armstrong, 2016):

1. Key mobilisers: active tweeters that got retweeted a lot.

2. Untwitting mobilisers: people who tweeted once or twice and got retweeted a lot

3. Moderate influencers, tweeters that got retweeted somewhat

4. Passionate participants, people who tweeted a lot but weren't retweeted as many times

These four categories do not seem very new. Every debate has members who contribute a lot and have a large influence. In Twitter storms, the key mobilisers are the most influential group while the passionate participants only drive up the number of tweets and have little influence on the debate. It does however prove that not all people online have the same influence in the debate, as is the aim of Twitter. The key mobilisers and the untwitting mobilisers are the people who are directly linked to the situation. In the #Ferguson storm the friends and family of the victim where retweeted the most, next to the media and the police reporting on the events that followed. So apart from the number of followers, the reach of one's tweet also depends on the connection people have with the Twitter storm.

Not only users determine what is seen on Twitter, algorithms play a huge part in what can be seen on someone's timeline as well. Algorithms are smart formulas that scan the content of tweets online in order to determine to what extend the must be made visible. They select what users will see these tweets and in what order (Gayo-Avello, 2013). There are several methods and codes to programme an algorithm. One of them is PageRank:

''PageRank algorithm aims to determine a numerical value for each document in the Web, such a value would indicate the ''relevance'' or ''authority'' of that given document. That value, also known as PageRank, spreads from document to document following the hyperlinks – previously it must be divided by the number of outgoing links. That way, heavily linked documents tend to have larger PageRank values, and those documents receiving few links from highly relevant documents (i.e. documents with large PageRank values) also tend to have large PageRank values.'' (Gayo-Avello, 2013)

This algorithm is the most used on social media and works in a similar way as Google does. Because it does not only look at the content but also at the relevance and the authority, it will likely exclude spam from being linked to. Relevance is determined by the words that are used in the content. Authority is determined by how many others have linked to the content, how many times it was seen by others and whether the agent who posted it is a real person or not. On Twitter a similar algorithm is used. This algorithm is called TwitterRank and also scores on relevance and authority. Next to selecting content on relevance it also looks to topic similarity between users (Weng, Lim, Jiang & He, 2010). This means that it will likely show content on topics users have tweeted about or are following now. For instance, one might see tweets from hospitals if one is reading a lot of content about diseases. No-one is controlling the algorithm, it is self-operating and once it is installed it becomes self-learning. The algorithm alone determines if content is made visible to a user.

In a Twitter storm, the algorithm will show the tweets that have the most retweets next to tweets that it expects to be the most interesting to the user. These tweets will therefore likely be ones by peers or like-minded people. Searching for hashtags and certain topics will help the algorithm to gain new content and learn more about what you might want to see (Gayo-Avello, 2013). TwitterRank is used to keep you on Twitter for as long as possible. This way the platform will make more money from advertisements. By showing things you might like, you will hopefully read on and spend more time on Twitter encountering more advertisements (Sriram, Fuhry, Demir, Ferhatosmanoglu & Demirbas, 2010).

2.5 Conclusion

So far, the conclusion is that the rule of law as a set of values protects citizens from tyranny and ensures that all have the right to a fair trial. When Twitter storms occur, they are mostly based on emotions. One emotion triggers another and reactions build on each other. People who are most directly involved with the object of the storm and people with the largest amount of followers, have the most influence. When Twitter storms occur, they can affect people from different rules of law. This thesis will therefore look at the values the rule of law upholds and not to individual nations' sets of laws. The values provided by Dicey (Jowell & Oliver, 2007) will be used in a comprehensive manner. Three elements will be reviewed: democracy, fairness and accountability.

3. Argumentation

In this section, the third and the fourth sub-question will be answered: How are individuals treated online by others and by the platform? And does this treatment do justice to the rule of law? As was explained in the previous section, these two questions will be answered by three values that can be abstracted from the rule of law. Firstly, democracy will be examined, secondly the value of fairness (right to a fair trial), and thirdly, accountability. These values capture the essence of both the history of the rule of law and the values formulated by Dicey (Jowell & Oliver, 2007). In all three sections, positive and negative effects are given. Each section ends with a concluding statement, summing up the main arguments given.

3.1 Democracy

Democracy is one of the most important values that the rule of law embodies. A state must fulfil several requirements in order to be a true democracy. In a functioning democracy, all individuals above a certain age can participate in and determine the political course of the state. In a working democracy, all voices are equal: everyone has the same chances to participate in public debate. It is the task of the government to provide this democracy. People should be able to inform themselves about different political parties and statements. The process should be transparent and accessible to all. Therefore, freedom of press and freedom of speech is needed. Without these two rights, the process will lose its integrity.

That is why in conventional trials, subjects touching on those freedoms will usually be handled with great care. This caution is not always exercised when Twitter and other social media replace official trials. When suspects are targeted online, the police can take action if there is reason to believe that a serious crime has been committed (Janoski-Haehlen, 2011). In Twitter storms where individuals are targeted and responded to by others this can lead to a direct form of 'punishment' or 'justice': the users decide what consequences are justified for what actions. It could be imaginable that in the future, trials with juries will be replaced with internet juries. Twitter can be the first step towards such as process where more people than the fifty or less jury members decide the outcome of a trial (Janoski-Haehlen, 2011).

In this sub-section, arguments for and against Twitter storms will be given that are related to democracy. These arguments review three features of democracy, namely: representation, safety and transparency.

3.1.1 Representation

As discussed earlier, democratic states have become too large for a system of direct representation. Models of indirect representation of the citizens were created (Bovens, 2003). The first argument in this section is about representation. As has been argued by internet optimists, the internet can enable more citizens to become politically active. This is also possible for social media like Twitter. It could be that Twitter is helping representation since it is a platform on which users can engage with each other and with politicians. New options for direct contact are given. Moreover, no matter one's social status, ethnicity or wealth, anyone can make a profile and send tweets. It might be so that Twitter is used to follow the news and inspiring persons, and can in that way help to form political opinions. Twitter storms form an extreme example of both of these possible positive contributions to representation. Twitter storms ask for awareness, they highlight certain topics for both the media and the government. Representation is thus enlarged as everyone with an account can vent their concerns and contribute. Questions should however be raised about the quality of this representation. In an offline democracy, people who devote their careers to political decision making or the prosecution of suspects are in charge of the discussion. They have a certain expertise and are trained to lead a fruitful discussion, using their know-how to come to a well-grounded opinion. Their time can be completely devoted to look for all the arguments and they have the ability to weigh and balance them. Society expects them to do this professionally and systems are in place to make sure they do so, instead of abusing the power invested in them by society. The media will for instance check if politicians speak the truth and act with integrity. What is seen as justice is decided by law and brought upon suspects by professionals with completed training. In order to become a judge, one has to follow a long trajectory of schooling and one must take an oath. When this oath is breached, one will never be able to operate as a judge again (Hague & Loader, 1999). This is different for the users that partake in Twitter storms.

First, Twitter users are not always held accountable for what they post online. It is not even always sure who owns an account and who is an authority online (Janoski-Haehlen, 2011). The responsibility to check if one is speaking the truth lies with the other Twitter users. If people in Twitter storms are indeed reaction out of emotion, it is questionable if they will check the facts behind the tweet or simply respond to it. One can easily follow others who are like minded, repeating their statements and therefore adding to the quantity of the discussion without controlling the quality of their arguments. People will not always scroll back to the start of the discussion and read what has been said originally. Due to the large number of tweets, this becomes impossible. Not only the number of tweets contributes to this non-controlling behaviour, due to the character of Twitter users, they will not spend much time on the platform. Since Twitter is mostly used for micro blogging and quick news updates (Kwak et al., 2010), people will spend only a short amount of time on the platform before moving to another activity online. This swift character will decrease the chance that people investigate what has truly been said online by others. They are likely to follow the information that is directly displayed on their timeline, and nothing beyond that. People will therefore first encounter the people they are already following. This may lead to a biased opinion since users of the platform mostly follow their friends and co-workers or the people they find inspiring and became 'friends' with online (Huberman, Romero & Wu, 2008). Online friends often 'help' each other in the discussion and enhance each other's statements. Facts become lost when people are copying their tweets. Only those who are truly passionate about a case will post multiple tweets about this case, or post opinions that differ from what has been written by others on their timeline (LeFebvre & Armstrong, 2016). So, in a Twitter storm, people are most likely to comment on others without validating them first. The tone of a tweet might also not always be clear. When people react to others they do not always know with which intent something was posted. Sarcasm, irony and scepticism might be overlooked. Differences in age, gender, culture, political opinion and religious beliefs determine how people use their language. This might lead to misunderstanding online when one does not ask for clarification and no visible emotions are displayed (Rao, Yarowsky & Shreevats, 2010). People from different backgrounds are therefore not always able to read between the lines or nuance of what has been said. You cannot ask for validation from others in Twitter storms. The Twitter storm #ShirtStorm contained 53,905 tweets (Morello, 2013). One will not look up all of the users who have sent those tweets. They simply reply on what has been said in the last few tweets and base their opinion on the discussion on those tweets.

Is it the responsibility of other users to check if what has been said before is based on facts? I am personally not sure if we can ask the users online to do so. People are on Twitter for a short amount of time, they mostly do so to vent their own thoughts in a micro blog or to stay up to date (Kwak et al., 2010). It might not interest them if others speak the truth online. Twitter storms hold an element of sensation, it can be amusing to follow these large discussions and participate in them. Checking the facts or write tweets that are based on more than only emotions might not be the most fun thing to do. It might also not lead to

gaining new followers or the most retweets. The amount of followers helps to determine how we behave online. While accounts used by professionals such as scientists or journalist might benefit from giving others the facts, other accounts might benefit more from making emotional statements and stirring up the discussion. Furthermore, not all users might have the know-how or the time to validate others.

Representation is not only about the validation of facts, it should also be of importance that all have an equal chance to participate and are heard by others and representatives. This has proven to be problematic online. Not all have access to the internet, let alone a Twitter account. When looking at the population of the USA for instance, it shows that they are not equally represented on Twitter (Mislove, Lehmann, Ahn, Onnela & Rosenquist, 2011). People who live in the countryside are less likely to have an account then those who live in cities. When looking at Twitter in age categories, people in their twenties form the largest group, followed by people between 30 and 40. Accordingly, elderly people are almost never active on the platform. Most of the Twitter users in the USA, who display gender on their profiles, seem to be male. These statistics found in the USA in 2009 seem to be still representative in 2018, and they are quite similar to worldwide statistics[2]. What's more is that people who have an account seem to usually be people with a higher education. People with no education or a practical education seem to be under-represented. So, the users of the platform are young people, with better changes in life (looking at capital). From these statistics, one can conclude that the Twitter population is a poor representation of the real population. What I personally find even more worrisome is that in Twitter storms the people that are most likely to be heard are the key mobilisers. These key mobilisers might be people who are directly involved or have a certain expertise. There is however also a chance that these key mobilisers are responding in a very emotional way and therefore become popular. Other users will be led by those statements. One can argue whether it is desirable if the most influential people are the ones who act out of emotion. Even though venting one's feeling from time to time might feel nice, it might not be the best way to contribute to an online discussion. Apart from content by key mobilisers, content that you are most likely to read is displayed. The algorithm will show you what it expects you will like the most. This can exclude you from content that will show you different opinions or arguments than you are used to read. Opinion forming via Twitter becomes limited and narrowed by these algorithms.

3.1.2. Safety

Another important condition for a functioning democracy is safety. Individuals must feel safe to express their opinions and vote. In a safe environment, discussion can flourish, helping all to come to political opinions and to openly and actively contribute in decision-making

---

[2] https://www.omnicoreagency.com/twitter-statistics/ Twitter demographics of 2018 by Omnicore.

processes (Bovens, 2003). Safety is also of importance when Twitter storms happen. It might help to create a safer environment if bad or criminal behaviour is pointed out online. This happened with Dusty the cat. In a YouTube video images were shown of a young male who abused his cat[3]. This cat was later called Dusty by people online. The video didn't land well and was seen as horrible misuse by its watchers. As a response to the abuse they formed an online search party. Within 24 hours, their online movement found the owner of the cat and the police was able to arrest him the next day. The use of the hashtag #AvanceMe and several images of the cat, helped to stop the abuse. A Twitter storm helped to stop injustice from happening. There are more cases imaginable were Twitter storms could indeed help to stop crime and make the internet and the world a safer place, using the voice of the people. It is however questionable if this is always the case. This argument will be explored in a later section. The use of a Twitter storm in order to do 'justice' must however happen in such a way that users can express their opinions without being threatened. Unfortunately, this is usually not the case in Twitter storms, in which especially women and minorities become targets online (Morello, 2015).

Internet trolls contribute to a poisoned environment online. The phenomenon of trolls has become relevant during 2017 because multiple scientists assume that trolls influence political campaigns and their effects. A troll is a person who enters online spaces and platforms to engage with others in a devious way. They can be compared to folklore characters that are called tricksters (Buckels, Trapnell & Paulhus, 2014). These modern-day villains want to upset others and stir up discussions for their own enjoyment. In the last five years, the amount of people acting as trolls online grew tremendously. A troll is usually not acting as a single account but creates fake accounts that operate on their own, retweeting the tweets of the troll. So-called bots help the troll to gain online status and credibility. It makes them look like real people. Because of these bots, trolls can gain influence. The more bots a troll owns, the more likely it is that he is influential in the debate since the number of retweets and followers determines his status on Twitter. The number of bots a troll owns can vary between 25 and over 25,000 (Buckels et al., 2014). Unfortunately, these individuals also engage in Twitter storms. The #shirtstorm as earlier presented became a movement that wanted to promote the rights of women. Individual women were targeted by trolls online after engaging in this discussion (Morello, 2015). For instance, Jacquelyn Gill, a female paleo-ecologist, received death and rape threats after she started to actively support the shirt-storm movement. Her colleagues wrote a computer programme that could find and filter these responses. They found out that a lot of them were sent by troll accounts. The Pew Research Centre for internet and technology is reviewing internet harassment every

---

[3]   http://knowyourmeme.com/memes/events/kenny-glenn-case-dusty-the-cat last visited on 20-06-18.

year in order to help the American government prevent and locate harassment[4]. Their last report in 2017 concluded that roughly four in ten Americans experience online harassment. This harassment could vary from offensive name-calling to physical threats to sexual harassment and even stalking. Cases of this mainly took place on social media like Twitter, since these are fertile spaces for trolls and other users who wish to hurt individuals because of the open character of the platform. People who suffered from violation felt less likely to engage in online debate again. Some even removed their accounts, suffering from real-life consequences such as fear for their safety or loss of social status. Status can be lost since victims of harassment might be blamed by others for the harassment. It could also be that these victims become shamed in or even excluded from a community, or they can be linked to sexual abuse. The death threats that were sent to Gill are no abnormality of Twitter misuse. Women participating in the #Me-too discussion that went viral in October last year were bullied and threatened online by users who said they stood up for the rights of men. They used troll accounts in order to massively spam these women and send hateful content to them (Morello, 2015). Anyone could make a troll account and start these ill motivated campaigns on Twitter.

Not only trolls make Twitter a bad climate for fruitful democratic discussions, anonymity can also lure people into online misbehaviour. The fact that you are sitting behind a screen and have no direct contact with others has proven to stimulate misbehaviour online (Lapidary-Lefler & Barak, 2012). This bad behaviour consists of calling other users names, threatening them and sexually insult or harm them (Dyer, Green, Pitts & Millward, 1995). In all, the behaviour seen in Twitter storms is caused by lack of personal contact with others. You do not see the impact of your actions, nor how others are hurt by your statements. Other people become dehumanised when their personal features and intonations are brought back to a picture and text, making it easy to call them names or worse (Lapidot-Lefler & Barak, 2012). Thus, when Twitter storms happen, people are tempted to respond in ways they would never respond to others face to face. Moving back to the case of Justine Sacco, where people responded that she should ''Lose her job and get AIDS'', this dehumanisation of individuals might clarify why people responded so harshly. They didn't have to worry about Justine's responses since those would never be more than texts typed on a screen. They probably wouldn't have made such comments in face to face contact, and if they would have, their effect wouldn't have been so devastating. According to Lapidot-Lefler & Barak (2012), it is especially the lack of eye contact that enables us to say hurtful things in Twitter storms or on online forums. Online anonymity stands for more than not only seeing the other. It also means that one can hide behind an account or use the same account with multiple people to bully someone. When others do not know who we are and will probably

---

[4] http://www.pewresearch.org/topics/social-media/2017/ link to the 20117 research report, last visited on 20-06-18.

never find out, we will also say and do things we wouldn't do in 'real-life' contact (Lapidot-Lefler & Barak, 2012).

It is partly the task of algorithms to ensure online safety. They should be programmed in such a way that they do not show harmful content or content that is posted by fake users. This content is labelled as spam and can be detected because no-one has linked to it or because certain words or topics are used in the content (Gayo-Avello, 2013). For example, the algorithm can be programmed for people who try to spread links to pornographic websites and ask users to validate their account if they are using Twitter as a means to promote these websites. They might be bots or trolls, linking people to viruses or illegal content. TwitterRanks is however very unable to protect the safety of the Twitter users (Gayo-Avello, 2013). Content that can be labelled as spam by humans can easily slip through the filters of the algorithm. This is possible because Twitter spammers are very diverse, the algorithm cannot keep up with the spammers and the ways they circumvent the terms of use and the filters.

Whereas Twitter storms can correct criminal behaviour if this is displayed online, they also form an opportunity for users to display this sort of behaviour. Twitter is not the safest environment to exercise your democratic rights and speak up for what you find important. When you do so, you risk your own safety, loss of status and might feel too threatened to leave your home. I therefore conclude that the gains from Twitter storms in terms of safety do not outweigh the risks people are exposed to if they engage in these discussions. The safety that a democracy needs in order to flourish is violated in Twitter storms. They can encourage people to act in morally wrong ways and hurt others for entertainment without facing the consequences of such behaviour.

3.1.3 Conclusion

Twitter storms can help to provide democracy for a rule of law or multiple rules of law since they form a way for people to vent their thoughts and raise awareness. Citizen initiatives could also happen online via these storms. Crime can be spotted online and stopped via a Twitter storm. But I personally do not think that these possible gains from Twitter use compare to the risk Twitter storms bring about. Because of the demographics of Twitter, online representation is not in line with offline representation. Groups are excluded from the discussion since not all people are on Twitter and not all users are equally displayed. Younger, highly educated people have more chances to participate than elderly people who are not on the platform. These statistics exclude Twitter as a means to represent the interests of the citizens. Moreover, Taking part in Twitter storms can harm one's safety. All can be at risk when venting their thoughts online. Emotional tweets can lead to Twitter storms that target an individual. Participating in these storms are bullies such as trolls or just anonymous users. They can limit others in their freedom by threatening them or encouraging certain behaviour. Discussions are not safe enough to be truly fruitful. Because of the lack of equal representation and questionable safety online, Twitter storms are

harming the value of democracy. Individuals are mistreated in these storms. Not only are they not safe, they become easily targeted when Twitter storms occur.

3.2 Fairness

This second section will evaluate arguments about fairness. Fairness is one of the values Dicey named (Jowell & Oliver, 2007). This means that all individuals living under the rule of law must have access to a fair trial. When they are prosecuted they should all be treated the same since they are all equal under the law. Gender, race or status should play no part in any trial and equality should be promoted. All trials must follow the same standard protocol. This ensures that the procedure is exactly the same for all, granting consistency. Even more, law should be known to all citizens in order to be fair. People cannot be punished for crimes that have not (yet) been described in the law or for acts that they could not have known to be illegal. This consistency not only holds for people who are prosecuted, but all citizens under a rule of law should be able to press charges against others and hold them accountable for their deeds. Whether the other is a king or a queen, a police officer or a judge, all can be put to trial by an individual if they are thought to have committed a crime. Lastly, all crimes should be handled in the same way. Persons who committed the same crime should suffer the same consequences. These consequences also are not dependent upon status or position, within the rule of law. So, if the penalty for theft was cutting of one hand, as it was in the Middle Ages, this should apply to you whether you are a nobleman or a peasant.

As was mentioned before, Twitter storms have the ability to point out potential danger and raise awareness for what users see as harmful behaviour. This section explores how Twitter storms affects individuals and if online the value of fairness is met. Fairness will be split into having access to a fair trial (procedural justice) and having the ability to press charges against others (responsibility).

3.2.1 Procedural justice

In order to decide if Twitter storms affect the value of a fair trial through procedural justice, it is necessary to assess what the concept of a fair trial is and what a fair trial entails. A study performed by Tom Tyler (1988) researched what the elements of fair trials were for suspects from the first contact with the police to the final judgement they reserved in court. He researched suspects and people who had contact with the police or experienced one or multiple trials. The interviewees were selected by a random sample. 47% of this group had called the police for help, 31% was stopped by the police and 23% had been to court. These different people were asked question about their contact with the authorities. Trials that were seen as fair by his research subjects had high ratings on the following six topics:

1. Control and representation: a suspect must be listened to and have the chance to explain their case to the authorities.

2. Consistency: trials should happen in the same way to all suspects and be in line with their expectations.

3. Impartiality: the people the suspects come across during their trial should be unbiased. They must treat them in an equal way, not favour others or let outcomes be influenced by personal characteristics of the suspects. There must be a neutral attitude towards the suspects.

4. Decision Accuracy: The authorities must have the right and enough information before coming to a decision about the suspect. Only with the right information and an unbiased look to all of the documents, a judgement can be made. A suspect is always innocent until their involvement has been proven.

5. Correctability: suspects must be informed about organisations where they can complain about the procedure of the trial and have access to these organisations.

6. Ethicality: the authorities must display behaviour of care towards their suspects, meaning that the show concern for their rights and stay polite (Tyler, 1988).

It is hard to translate this study to the Twitter storms that occur now, for it is not clear who has the authority in these storms. This seems to be the company Twitter, but they describe in their policy that they will not actively look for misuse of the platform. If posts do not contain certain images (nudes or porn) Twitter will allow them on their platform until abuse is reported by users. This makes the users responsible for what is posted online and what is seen as offensive. Twitter storms are the product of people responding to one's statements. The people who are retweeted the most or whose tweets have the greatest impact determine the course of the discussion. However, it is however, mostly the direct environment which takes steps towards 'real-life' consequences that affect the individual. In some cases, the police can step in and arrest the individual. The executive power is on Twitter as well, looking for criminal behaviour and terrorist threats. Besides the interference of the police, people who feel violated online can sometimes call the authorities and report misbehaviour. The police or the public prosecutor may then file the case. Arguments about claiming your legal rights will be made later. In most Twitter storms there is no authority that can be evaluated with these features of a fair trial. However, these features can be applied directly to a Twitter storm.

In the section on democracy, it was argued that not all people online are heard equally well. Trolls and algorithm randomness can influence the aim of Twitter to give everyone an equal chance to vent their feelings. Tweets with strong emotions drive the storm. Individuals who are blamed will therefore not have the chance to explain their case to other participants. One cannot defend oneself when one's status is being harmed online. Before individuals are interviewed by the media or make their statement known, they often have already lost their status and suffer the consequences of what they are blamed for. When the presenter of MTV's show Catfish, Nev Schulman, was accused of sexual abuse by one of the girls in his

show, a Twitter storm broke out and all of his known missteps in the past were vented online[5]. He later claimed that he did not approach the girl and that several witnesses could confirm his behaviour towards her, but MTV still had to stop the show. His status was lost and the network no longer wanted to be associated with the presenter. Thus, before the police could investigate whether Schulman had behaved inappropriately, he had already been accused and had suffered loss. He could not defend himself; his statements were not heard by the people participating in the storm.

Consistency is met in the Twitter storms, since they seem to be following the same process. Individuals who become suspects of immoral behaviour can expect the same treatment online by both the algorithm regulating the platform and the curve of other users response (Morello, 2015). There will be people calling them names, bullying them or even worse. Other will stand up for the individual and try to hush the arguments of the opposite side of the discussing. The storms are temporary and will, luckily, not stay viral. It is nonetheless questionable if this makes them any fairer. Twitter storms are driven by people, trolls and a certain randomness caused by the algorithm. Participants may act out of their emotions. They are certainly not unbiased towards the 'suspect' of an immoral action. Since they can act anonymously they might target people based on their ethnicity, gender or other personal features. Most of the people online do not have a neutral attitude towards other users (Morello, 2015). Next to a lack of impartiality there is no decision accuracy online. Due to the size of the debate, people will not read all the comments or the original post. Tweets are only 280 characters long and can therefore not give enough information to come to a fair judgement. Evidence is missing and the other side of the debate is not heard.

However, there are organisations where individuals can complain about the Twitter storms. One of these organisations is Twitter itself. In Twitter's policy, a guideline is provided for when abuse happens on their platform. Twitter will take steps if users break the terms of use they have agreed to when they signed up for the platform. Users might be banned from the platform or single tweets will be removed. In serious cases, the policy also advises individuals to contact the authorities. Twitter might mean well by sending people to the authorities and have users banned or tweets removed. This does however not solve the problem Twitter storms are causing. The size of a Twitter storm is problematic, and so it is impossible to notify Twitter about all the people who respond negatively to your content. Even more, not all these negative responders violate the terms of use. Most of them do not, since they are only venting thoughts and not using curse words or violent images. People are allowed to a certain level of negativity online towards others. On top of that, Twitter is not able to block a person from making a new account when his old one is banished from the platform. Therefore reporting abuse to Twitter might not always satisfy the targets of Twitter storms. The same can be said for reporting a Twitter storm to the authorities. The

---

[5]   https://www.theverge.com/2018/5/17/17365764/catfish-host-nev-schulman-sexual-misconduct, lastly visited on 20-6-18.

police are not able to remove tweets from other users, finding them and talking to them about their misbehaviour is a costly and time consuming process. The same goes for trials, while waiting for a response from the authorities, the Twitter storm might have already passed. Both the executive and the legislative power cannot keep up with the speed of the internet (Bovens, 2003). The rule of law might therefore be harmed when Twitter storms happen. The authorities cannot stop them and are powerless against the results of such a storm.

Arguments for ethicality, the display of behaviour of care, online are in line with the arguments for being unbiased towards agents or decision accuracy. People behind computer screens are easily tempted into showing their worst behaviour since it is not visible that they are behind their actions. The rights of other persons can be violated and they certainly are not respected actively online.

When looking at Twitter storms as if they were trials, they would absolutely not qualify for fair trials. The features of a fair trial are not at work online, since agents in Twitter storms aren't treated with respect. If Twitter storms happen and targeted individuals lose status and suffer real-life consequences before an official trial can happen, this makes them a violation of the rule of law.

3.2.2 Responsibility

A different argument can be made for responsibility. Responsibility means that individuals can prosecute others if they acted immorally. This enhances the rule of law since all individuals must follow the law and should be held accountable for their actions. Persons must be able to report people who violated their rights and governmental organisations that do not follow the constitution. Twitter claims that they can help to call others out when they acted wrongly. Twitter storms can draw the attention of the media and the authorities. Especially when an individual is harmed by a company, social media can help to bring justice. Since the last decade, companies have started to invest more in social media. Arranging their customer service via these online platforms and using them for online marketing and advertising (Jansen, Zhang, Sobel & Chowdury, 2009). The products these companies create, become popular through online advertising, using hashtags and ratings of different users. However, entering online platforms also poses a threat to companies. When they fail to deliver, individuals can use Twitter, and Twitter storms, to complain and make sure that they get what they paid for. Complaining over Twitter has three advantages. First of all, your complaint reaches a large amount of people in a short amount of time. Secondly, all can see your post at the page of the company. This helps to put pressure on the company to act and respond to your complaint. If they choose to ignore it, they may damage their status since other Twitter users see that they did not take one of their customers seriously. Thirdly, complaining over Twitter causing a Twitter storm can unite individuals fighting for the same cause. Individuals that were first helpless against huge companies like Starbucks or McDonald's can now unite and enforce better treatment of customers.

One can however ask if this is a just way of Twitter use. Companies have to act in order to maintain their status as a friendly company or a company that cares about its consumers. It becomes very easy for consumers to claim that they have been treated wrongly by the brand or that there is something wrong with their product. While 280 characters can easily be used to say that something is wrong, they are way too few to investigate what the actual problem is. When people cause Twitter storms online in order to get a refund or a new product, the company must always give in to these individual claims. There is no chance to prove that the defect is a fault of the consumer.

Not only companies can be hurt by Twitter storms, Twitter users can also exercise political power, according to Shirky (2011). This happened for instance in January 2001:

''During the impeachment trial of Philippine President Joseph Estrada, loyalists in the Philippine Congress voted to set aside key evidence against him. Less than two hours after the decision was announced, thousands of Filipinos, angry that their corrupt president might be let off the hook, converged on Epifanio de Los Santos Avenue, a major crossroads in Manila. The protest was arranged, in part, by forwarded text messages reading, "Go 2 edsa. Wear blk." The crowd quickly swelled, and in the next few days, over a million people arrived, choking traffic in down town Manila. The public's ability to coordinate such a massive and rapid response - close to seven million text messages were sent that week - so alarmed the country's legislators that they reversed course and allowed the evidence to be presented. Estrada's fate was sealed; by January 20, he was gone. The event marked the first time that social media had helped force out a national leader. Estrada himself blamed "the text messaging generation" for his downfall.''

These, and later examples, show that Twitter storms can help citizens to arm themselves against their government if they feel it violates their rights. Protest movements can be enabled because of Twitter. Mass movements can gather and inspire others to raise their voice against oppression. This might, such as in the case of the Philippines, not always lead to a better government, but it will create feelings of oneness and solidarity (Shirky, 2011). Twitter storms could in that way enrich the rule of law. As was earlier described with the case of Dusty the cat, the same goes for individuals. The local authorities are often not capable of overseeing all the crimes that are happening online or in their neighbourhood offline. The internet and online initiative can help them to make society safer for all (Bovens, 2003).

3.2.3 Conclusion

The value of fairness consists of two main elements, access to a fair trial and responsibility. As was proven by reviewing criteria for a fair trial, Twitter storms do not provide the environment for fair judgement. There is no authority and people online are at least as biased as they are offline. They seem to be acting out of their emotions, treating 'suspects' without respect for their rights. What is said online is mostly harsh and out of proportion. As

long as the terms of use of Twitter are met, people can say whatever they please. Due to the size of the storm no-one can oversee what is happening and who is responsible for it. The right to defend yourself is also lost in the size of the storm. Individuals are unable to make their case to the other participants. This makes the storms unfair. Moreover, people cannot expect the same outcome of a storm as consequences are unpredictable. The variety of outcomes proves also that Twitter storms cannot be called fair.

However, the rule of law is also strengthened by Twitter storms since they have the ability to arm the masses against political abuse or big companies. Twitter storms have the potential to raise awareness for crimes hold others responsible for their actions. These two advantages for the rule of law should be taken into account. Nonetheless, in terms of fairness Twitter is not making the society fairer for individuals.

3.3 Accountability

The last arguments that will be discussed on how Twitter storms affect the rule of law are about accountability. Accountability in a legal context means that rules and laws should be made by the citizens living under them (Jowell & Oliver, 2007). This justifies that those citizens can be punished when they do not follow the law. Aside from enabling authorities to exercise their power, it makes sure that the law benefits all of the people living under the law. Or that there is at least some consensus about the law. Much has been said about Twitter storms and whether they help to hold individuals accountable for their actions. However, the question remains how Twitter storms affect rulemaking.

3.3.1 Web 2.0

As Bovens (2003) predicted, the internet is being used to involve more people in political decision making. Local governments go online and let citizens participate in decision making, shifting power from the government and bureaucratic structures to the people. This process is called horizontalisation. The USA is one of the countries where the internet has been used recently to include more people in policy making. Under Obama's government, Web 2.0 was created in order to sustain e-rulemaking. E-rulemaking is a process in which citizens can actively contribute to new policies, and read plans of the government and their motivations (Farina, Newhart, Cardie & Cosley, 2011). The American government hopes to be more transparent, achieve greater participation and collaboration. Next to the three classic powers in the rule of law (legislative power, executive power and judicial power), in the last decades a fourth power gained influence. This fourth power consists of agencies and officials that make rules and decide how new laws should be implemented. The fourth power is not chosen by the citizens and is sometimes therefore seen as undemocratic (Bovens, 2003). This fourth power should therefore become more transparent and collaborate closer with citizens (Bovens, 2003; Farina et al., 2011). Web 2.0 gives citizens the opportunity to comment on new rules through the use of the internet. When new rules are being made at the state level, citizens online can read the new protocol and respond to it by leaving comments. Discussions between participants are supported and their outcomes can

influence the protocol. In this way, new rules would become more democratic and the value of accountability would be met. By allowing people to read the content of rules before they go into effect, the government hopes to become more transparent. However, this has proven problematic. Many people are oblivious of the possibilities and the new policies that are being designed. Web 2.0 has not yet reached the popularity the government would like it to gain. In order to solve this issue, the Regulation Room was created. This control room of the project aims towards promoting the project. They do so by informing the media, and one of the latest steps to increase attentiveness was the use of social media. Via platforms like Twitter and Facebook, citizens can receive notifications when new protocols are posted online. The Regulation Room hopes that they will let the people they are connect to by these media now that they can also respond. This new formula of citizen initiative is still being tested and researched in order to improve user-friendliness and effectiveness. Twitter is used most since it is the easiest to track and investigate, because other social platform are often more private. For example, many Facebook posts are only visible if people share a connection online, such as Facebook friendships or liked pages (Farina et al., 2011).

Another problem, next to the low response rate, is the fact that by using social media the government allows platforms like Facebook and Twitter access to their messages, statistics and more private information. Next to the fact that they have to share their statistics, the algorithms of the sites decide who is seeing what. Not all citizens will therefore see the content. The algorithm will show the content to those it might be of interest to. People with political interests who follow the Regulation Room or similar institutions will have a better chance at responding on the content than those who have no internet linkage to politics. This strengthens the representative problems mentioned earlier. These algorithms might also make it harder to investigate who is seeing the content and who isn't, since they evolve and decide for themselves what is of interest to whom.

3.3.2 Raising awareness

Twitter storms could help the Regulation Room as they indicate stakeholders. These stakeholders can then be addressed directly by members of the Regulation Room. This happened for instance when a new rule on texting while driving was being implemented. Researchers wanted to see whether responding to people using hashtags about driving and texting would help to increase participation. They looked for Twitter storms that were the product of traffic accidents in which the suspects used their mobile phones while driving. In these storms, users online vented their concerns about driving and the use of mobile devices. They were easily traceable because they used hashtags like #donttextndrive. Members from the Regulation Room traced these (concerned) users and sent them an invitation to participate in the discussion around a prohibition on texting while driving. They expected that these people wanted to express their concerns and contribute to a new law. Users who were concerned with traffic safety could become stakeholders in the discussion. Unfortunately this didn't prove to be very successful. Most of the people who were

addressed did not respond and did not subscribe to Web 2.0 (Farina et al., 2010). Next to the use of Twitter storms, promoting Web 2.0 and its topics on the Twitter page of the Regulation Room did also not increase responds in the Web 2.0 environment.

Responding to Twitter storms might not have been the best way to increase citizen participation in rulemaking. This might however be different if Twitter storms start after the new policies are promoted in news media. When television (CNN) and newspapers (Washington Post) are used, responses increase. Using the popularity and the reach of this media, more and more people will become aware of what is going on in Web 2.0 and add to the discussion. This can create Twitter storms if the opponents of such new rules start to tweet. They will attack the government and the law itself. These emotional comments reach a lot of people if they go viral. This might help to inspire others to also vent their thoughts. If these people can be moved from Twitter to Web 2.0, their input can be used. The news platforms play a crucial role in this. If they provide information and links to Web 2.0 and the topic, more people will sign up and use their political power. If the news sites do however not include direct links, discussion will usually only take place on social media (Farina et al., 2010).

From this research can be concluded that it is hard to motivate people to contribute to online rulemaking in Web 2.0. This might be the same for similar initiatives. More research on this is needed however. Social media can help to decrease unawareness, but they will not always drive people to take the time to actually contribute to better laws. What is worrisome is that social media algorithms will decide who sees the content that can be responded to. This creates unequal chances to contribute. Causing Twitter storms in order to gain response in Web 2.0 might also come with some unwanted side effects. Twitter storms cannot be controlled or completely overseen. This means that the people who will come from social media to the Web 2.0 environment might be biased by misinformation or manipulated by what they saw on their timelines.

3.3.3 Quality of rulemaking

Twitter storms might benefit political participation of citizens and the creating of new democratic rules. However, sceptics of the use of the internet argued that most of the attempts of citizens to contribute would be of concerning low quality (Bovens, 2003). Participants would be badly informed, ill-motivated and might limit others from participation. The Web 2.0 project evaluated how the use of Twitter to contribute people in decision making would affect the quality of debate in the Web 2.0 Environment. They found that participation can be expected of four different groups of people: the good (1), the bad (2), the ugly (3) and the unknown (4) (Farina et al., 2010). Contributions of these groups will be evaluated briefly.

The first group, marked as ''the good'' by the researcher of the Regulation Room, are the people who are generally concerned about the new law. They take the time and effort to

read the proposals and comments that have already been posted online and then try to give their own opinion. They write typically long statements about their view with dense arguments. Others respond to these and dialogue is started. These people are often direct stakeholders and have experience with the topic. They are valuable for rulemaking since they can offer their knowledge. However it questionable if this group forms a sound representation of the population. Chances are that due to their experience and interest in these topics they are not. It would be best for online rulemaking if all participants showed the behaviour of the good. The government can however not demand this and it is debatable if they should. For example, if they tried to do so by installing a certain threshold for participation, many people would be excluded.

The second group, ''the bad'', are people who are seduced by the design of the American method of participation and what to vote. The design of rules and protocols makes them want to immediately give their opinion. They often do not read all the texts since these texts are long and have a high density. They often cite studies and can therefore be said to have an advanced reading level. Policy makers can try to alter their language and design but this does not affect the 'bad' responders:

''recognising its appeal, while being sceptical that it will work. The voting instinct may be so strong (particularly in the context of public participation in government decision-making on a Web 2.0 site) that any accommodation will sabotage efforts to create a new participatory culture that makes higher demands on online community members '' (Farina et al., 2011).

Therefore, the government should take into account that even when there was a high representation on Web 2.0, the representation might contain people who do not understand what is going on.

The third group of people, ''the ugly'', also decrease the quality of the discussion. These people show the same behaviour that is displayed in Twitter storms. Due to the absence of real-life personal contact, they (intentionally) hurt others or ridicule the law proposals. Even though the Regulation Room formulated rules for participation such as "Comment on content, not on the contributor'', most of the people responding in the case of texting and driving did in fact comment on contributors. It is hard to stop such behaviour in Web 2.0 because it is not always clear if people are simply using their right to free speech, or abusing it.

Lastly there is the fourth group, ''the unknown''. These people are described as lurkers, people who only view the discussion but do not contribute. There might be different reasons to not contribute to a discussion: people might feel sceptical towards the government, not understand the content or feel that others have already posted their concerns. The Regulation Room strives to lure these lurkers into contributing even though they might not feel like they should. A large part of people that will only read the content will however

remain inactive. It is unclear how a government should deal with this form of participation (Farina et al., 2011).

3.3.4 Conclusion

Twitter storms can help to raise awareness for citizen participation in rulemaking. If the government responds to these storms, people can add issues to the political agenda. When projects like the American Web 2.0 enable citizens to contribute to new rules, social media can help to engage more people. This will lead to a more transparent government, more participation and collaboration of citizens and therefore an enhanced rule of law. However, the quality of this form of citizen participation might be questionable. Not all participants behave as projects like Web 2.0 would want them to. They enable room for discussion but this space can become filled with personal attacks and other hateful comments towards both others and the government. Freedom of speech and respect for other people are collapsing on these platforms, making it hard for governmental organisations to stop those who have ill motivations. Platforms like Web 2.0 can become the new target of trolls. Just as in Twitter storms, people can be easily harmed or threatened. The rule of law is however enhanced by projects such as Web 2.0. These projects form a positive promise towards our future where people are actively contributing to new rulemaking, shifting power from the government to individuals.

4. Conclusion

Now that all arguments concerning the values of rule of law have been reviewed, it is time to make up the balance. In this section, the main question of the thesis is answered: *How are Twitter storms influencing the ideal of a rule of law?* This is done by a summary of the arguments that have been made. After this recap, the question is answered, weaker points of the study are named and recommendations for future research are made.

The rule of law has been described as a system of values which can be enhanced or violated via the internet. Law philosophers, such as Bovens, predicted that the internet could be a promising new medium to include more people in public decision making, as it is easily accessible for all and reduces the space between rule makers and the public (Bovens, 2003). However, the internet can have some disadvantages as well. I have argued that Twitter storms form one of these disadvantages. They form a danger to different aspects of the rule of law. One of these aspects is democracy. When Twitter storms occur they seem to enlarge public debate and let all contribute in 'discussions' online. However, it is arguable whether a real discussion is taking place. Users are adding to what has been said before them, likely without reading the original content of the tweets that people reacted to. Algorithms display what they think would most likely interesting or likeable for the users in order to keep them on Twitter for as long as possible. This will lead to validation of one's believes instead of

showing people new opinions and arguments. Users will respond to one another without the know-how needed for a fruitful discussion. This makes storms nothing more than avalanches of thoughts. Even more, the people that have the most influence in these storms are those who respond out of emotion. Lack of expertise and time to investigate statements disqualifies most Twitter users from having founded political discussions when Twitter storms happen. What is also of concern is the fact that citizens are not represented equally on the platform. Twitter consists mostly of people in their twenties who have received a higher education. Elderly people and those living in the countryside are under-represented. The lack of safety also forms an obstacle for enhancing the rule of law. In order for a rule of law to flourish, citizens must be able to speak their minds without being threatened. While algorithms should protects users on Twitter from spam, the platform is still used by trolls and others who are ill motivated and wish to shame or threaten others for their own entertainment. This, and the absence of personal contact online, makes Twitter an unsafe environment to engage equally with each other. People become dehumanised and in Twitter storms we can easily violate others since we do not have to deal with them face to face. The consequences of our contributions to the storm will play out where we cannot see them.

The second feature of the rule of law that was discussed was fairness. All should have the right to a fair trial and should be able to hold others accountable for their actions online. Twitter storms could be a step towards online juries and can enrich the rule of law since they can point out possible crime. This is nonetheless undesirable as Twitter storms do not contain the same elements as offline trials that are seen as fair display. Suspects get shamed for the entertainment of others, leaving them with the loss of social status and other damage that might be far out of proportion to their perceived crimes. If they prove to be innocent, this loss cannot be repaired. The authorities are powerless against the size of most Twitter storms. They cannot be overseen, let alone stopped. Yet not everything about Twitter storms is negative. The might help individuals and groups to stand up to large institutions such as corporations or governments. Movements can be formed online and used in offline protests. Awareness can be asked for those who feel powerless against bureaucratic systems. Twitter storms do attract news media and tablets, and when they report on the Twitter storm, actual change can be accomplished.

The third and last feature of a rule of law is accountability. Citizens must be able to make or contribute to the rules the live under. Due to horizontalisation, the internet can enable more people to contribute to rulemaking. However, this so called e-ruling has proven to be hard to organise satisfactorily. People might not want to engage in rule making and if they do, the quality of this representation might be questionable. Not everyone carefully reads the contents of new legislations, and not everyone has the right motivations when visiting online environments to participate. Twitter storms were investigated to see if they can increase participation in e-rulemaking. This has shown to only be an asset for online rulemaking if the discussion is reported in news media. Due to the low quality of rule-making and the lack of contribution, the way we use e-rulemaking and Twitter storms now does not seem to be

very fruitful for the rule of law. More research to decrease the gap between citizens and their government is needed.

Twitter storms harm individuals. They are treated poorly online by other users, especially if they become the victims of storms. The same goes for the features of the rule of law. Twitter storms do not do justice to the rule of law. Values such as democracy, fairness and accountability are not enhanced online in these storms. Hence, the ideal rule of law is affected by Twitter storms in a mostly negative matter.

There are however some weak sides to this thesis that should be taken into consideration. For instance, the work that was used to describe the relation of the internet and the rule of law by Bovens stems from 2003. It is not clear whether more recent findings argue for other connections between the rule of law and online media such as Twitter. Social media like Facebook and Twitter did not yet exist in 2003. Therefore, their influence cannot be fully taken into account by Bovens (2003). Twitter did not exist in 2003 and the role of social media was smaller at the start of this century. Since the internet and social media are both rapidly changing phenomena, there is a realistic chance that there are new effects on the rule of law that should be taken into account. Twitter is a young platform and is therefore still evolving. While several studies on the features and demographics of Twitter were used in this thesis, not all is known about this social medium. Important advantages or disadvantages to its users or to society might have been overlooked.

Next to this newness of the medium and the somewhat outdated influences of the internet on the rule of law, more research should be done to the role that algorithms play in Twitter storms. This thesis touches upon a few known effects of the algorithms used on Twitter, there is however no material available on the role of algorithms in actual Twitter storms. It might be that algorithms behave differently during Twitter storms and display other content or review tweets differently. This is still in the dark. Furthermore, many of the examples, such as Web 2.0 and tele-voting, are situated in the USA. The USA might be a clear example of a rule of law since the three powers are clearly separated. It is however questionable if these examples can be translated to other countries. The Obama government has put a lot of money and effort into enhancing the rule of law through internet use. This is not the case in many other countries with a rule of law. Only recently the British government started to experiment with similar e-rulemaking procedures. More research towards e-rulemaking and the relation to e-rulemaking and Twitter storms is needed. Only then we can definitively conclude whether Twitter storms are usable for the enhancement of accountability.

5. Literature

Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, *24*(3), 233-239.

Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of marketing research*, *49*(2), 192-205.

Bovens, M. A. P. (2003). *De Digitale Republiek: Democratie en rechtsstaat in de informatiemaatschappij*. Amsterdam University Press.

Bovens, M., & Loos, E. (2002). The digital constitutional state: democracy and law in the information society. *Information Polity*, *7*(4), 185-197.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and individual Differences*, *67*, 97-102.

Crawford, C. S., Mack, N., Eugene, W. & Gilbert, J. E. (2015, August). Televoting: Secure, Overseas Voting. In: *International Conference on Human Aspects of Information Security, Privacy, and Trust* (pp. 487-494). Springer, Cham.

Dyer, R., Green, R., Pitts, M., & Millward, G. (1995). What's the flaming problem? or computer mediated communication–deindividuating or disinhibiting? In: M. A. R. Kirby, A. J. Dix, & J. E. Finlay (Eds.), *People and computers X* (pp. 289–301). Cambridge, UK: Cambridge University Press R. Kirby, A. J. Dix, & J. E. Finlay (Eds.), *People and computers X* (pp. 289–301). Cambridge, UK: Cambridge University Press

Farina, C. R., Newhart, M. J., Cardie, C., & Cosley, D. (2011). Rulemaking 2.0. *U. Miami L. Rev.*, *65*, 395.

Gayo-Avello, D. (2013). Nepotistic relationships in twitter and their impact on rank prestige algorithms. Information Processing & Management, 49(6), 1250-1280.

Hague, B. N., & Loader, B. (Eds.). (1999). *Digital democracy: Discourse and decision making in the information age*. Psychology Press.

Huberman, B. A., Romero, D. M., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. *arXiv preprint arXiv:0812.1045*.

Janoski-Haehlen, E. M. (2011). The Courts are All a Twitter: The Implications of Social Media Use in the Courts. Val. UL Rev., 46, 43.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the Association for Information Science and Technology*, *60*(11), 2169-2188.

Jowell, J., & Oliver, D. (Eds.). (2007). *The changing constitution*. Oxford University Press.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600). ACM.

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in human behavior*, *28*(2), 434-443.

LeFebvre, R. K., & Armstrong, C. (2016). Grievance-based social movement mobilization in the# Ferguson Twitter storm. *New Media & Society*.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., ... & Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review*, *39*(5), 22-31.

Mislove, A., Lehmann, S., Ahn, Y. Y., Onnela, J. P., & Rosenquist, J. N. (2011). Understanding the Demographics of Twitter Users. *ICWSM*, *11*(5th), 25.

Morello, L. (2015). Science and sexism: In the eye of the Twitterstorm. *Nature News*, *527*(7577), 148.

Norris, P. (2001). *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge University Press.

Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications*, *20*(1-2), 117-128.

Pitkin, H. F. (2004). Representation and democracy: uneasy alliance. *Scandinavian Political Studies*, *27*(3), 335-342.

Rao, D., Yarowsky, D., Shreevats, A., & Gupta, M. (2010, October). Classifying latent user attributes in twitter. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents* (pp. 37-44). ACM.

Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign affairs*, 28-41.

Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., & Demirbas, M. (2010, July). Short text classification in twitter to improve information filtering. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (pp. 841-842). ACM.

Tamanaha, B. Z. (2004). *On the rule of law: History, politics, theory*. Cambridge University Press.

Trottier, D., & Fuchs, C. (Eds.). (2014). *Social media, politics and the state: Protests, revolutions, riots, crime and policing in the age of facebook, twitter and youtube* (Vol. 16). Routledge.

Tyler, T. R. (1988). What is procedural justice?: Criteria used by citizens to assess the fairness of legal procedures. *Law and Society Review*, 103-135.

Weng, J., et al. (2010). TwitterRank: Finding topic-sensitive influential Twitterers. In Proceedings of the third ACM international conference on web search and data mining (pp. 261–270).