Bachelor thesis
Computing Science

Radboud University

Mutual privacy and good manners on the social web

Author:
Onno de Gouw
s1025613

First supervisor/assessor:
dr. ir. Eelco Herder
eelcoherder@cs.ru.nl

Second assessor:
dr. Hanna Schraffenberger
hanna.schraffenberger@ru.nl

August 23, 2021

Abstract

On social media we read comments of our family, friends, acquaintances and even strangers. These comments are sometimes positive and fun to read, but sometimes they might be offensive, threatening or even breaching the privacy of other people.
Within this research paper, we explore implicit social norms, behaviours and privacy expectations on social networking services by trying to identify to what extent people intervene online 'friends' that behave undesirable and breach other people's privacy. Also, we explore the bystander effect and to what extent it influences whether people actually intervene or not. Finally, we describe a possible and very general design concept for social media, with the ultimate goal of increasing the degree to which users of social media intervene and stimulating good manners on the social web. Our research is a qualitative exploratory research in which we have conducted in-depth interviews among users that intensively make use of social media.
Within our research, we find that deciding to intervene on social media is something very personal, as each individual has its own implicit norms and social expectations. People often intervene their online 'friends' using a private approach. Additionally, we find that the severity of a privacy violation and the relationship one has with a privacy violator are two important factors influencing whether someone chooses to intervene.
Besides this, we also find that the bystander effect in the way it is presented in related literature is quite limited when it comes to intervention on social networking sites. However, users do adapt their behaviour and use of language if they are aware of the presence of 'online bystanders'.

Contents

Chapter 1

Introduction

1.1 Problem formulation

Many people make use of the internet in different ways. One of these ways is by being active on social networking platforms where you can build social networks and relationships with other people online. There also exist many different social media services and websites that people use, like for example Facebook, WhatsApp, TikTok, Snapchat and Twitter. On these different social networking sites, we all read comments of our family, friends, acquaintences and even people that we only know vaguely or that we do not know at all. These comments can be positive and fun to read, but sometimes they might also be very negative, offensive or even threatening. In the latter case, we could also speak of undesirable behaviour. We generally don't like it if someone breaches our privacy or comments a very rude text underneath a picture that we have posted on our profile. However, just like there are norms and values in our offline society, there are also norms and values that we apply online. It might be difficult for people to determine the intention of a certain post or comment placed by others, because to one a criticizing comment might be rude, while to someone else it might simply be a way to provide their feedback or their opinion on something. Therefore, it might also be difficult for someone to determine whether they should intervene if they see someone they know behaving undesirable on a social networking platform, or whether they should leave it. This brings us to our research, in which we try to investigate where this boundary is for people and what influences it when it comes to intervening or not intervening if someone on social media acts undesirable. In particular, this research focuses itself on a specific type of undesirable behavior that occurs on social networking sites, namely breaching other people's privacy.
Different people might have different expectations about when, whether and how you should react on someone else if they act undesirable and breach others' privacy. These different expectations can occur because of different

factors. For example, the type and strength of the bond you have with the person that performs the undesirable behaviour could be of influence and also the norms and values about when something is considered a breach of someone's privacy might differ from person to person. Next to this, also social psychological effects like the so-called bystander effect might play a role. These different factors are all to be explored.

1.2 Research question

In order to further investigate this boundary between intervening and not intervening on social media that we discussed earlier, we have composed the following research question which drives our research and where we try to find an answer for:

"To what extent do people intervene their online 'friends' on social media, behaving undesirable by breaching other people's privacy and sharing content of others without permission."

As explained before, we focus ourselves on one specific type of undesirable behaviour in order to narrow down the scope of our research; breaching other people's privacy.
Next to this main research question, we will also focus ourselves on researching whether the social psychological bystander effect, which basically states that individuals are less likely to offer help to a victim when there are other people present [25], could play a role in these situations as well. Therefore, we have defined the following sub-question:

"To what extent does the bystander effect influence whether people intervene their online 'friends' on social media, behaving undesirable by breaching other people's privacy and sharing content of others without permission."

1.3 Motivation

There are multiple reasons why our research is important and why answers to the research question are relevant and add to the body of knowledge in human-computer interaction and societal aspects of security and privacy. First off, the amount of users that makes use of social media is still growing very rapidly and will continue to grow. According to the statistics of Statista, there will be over 4.40 billion social network users worldwide in 2025 [22]. To show the difference: in 2021, only 3.78 billion users made use of social networking services. This can also be seen in Figure 1.1:

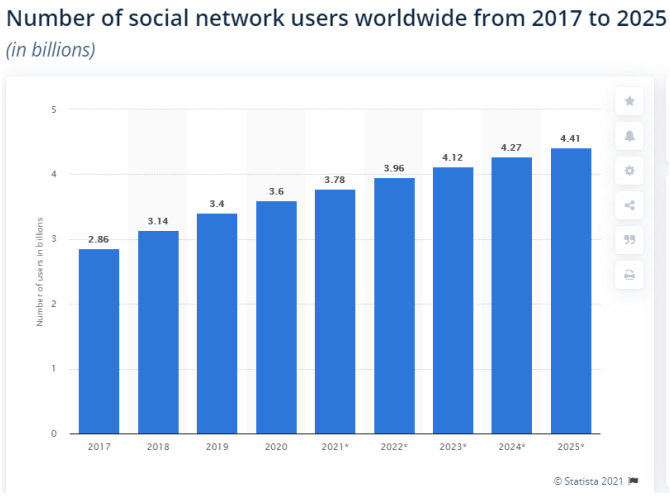

Figure 1.1: *Number of social network users worldwide from 2017 to 2025.*

Because of the fact that these numbers are still increasing, the amount of comments and posts will also grow and therefore the amount of privacy breaches and potential interventions will also grow. This shows that our research is currently relevant and will still be relevant in the future as well.

Secondly, not much is known about why people sometimes choose to intervene if they see something happening on the social web that is undesirable and why they sometimes choose to let it go and do not intervene. There is however, an extensive collection of literature already that relates to this topic. Much research focuses itself on topics like why people unfriend each other [23], different sanctions people apply on social media [18], different types of behaviour that occur [11], contextual privacy [16], fake news and peoples' reasons to share or to not share certain personal information online [28]. For example, there are lists of types of behaviour and possible sanctions that occur on social media and they show what behaviour people find undesirable and what not, but they leave questions about why and in which cases they will react on this behaviour, and what influencing factors there are, unanswered.
This in total makes that it is meaningful to research this boundary for people when it comes to intervening or not intervening when someone acts undesirable in the online context.

In this research paper, we will first introduce the related work in chapter 2. In chapter 3, we describe the research that has been carried out. It will contain the methodology we applied, the results we found and also a discussion. Finally, we will draw conclusions from our findings in chapter 4 and give some concluding remarks.

Chapter 2

Related Work

Before we present the actual research and the methods used together with the results generated from it, it is important that we describe how this research relates to existing theory. We will review relevant literature and relate our own research to specific knowledge that has already been established in other research projects.

2.1 Online behavioural norms

Just like there exist different social norms and values between people when they interact with each other physically outside, such norms also exist in the online world. Each social media user has their own social norms, defining what kind of behaviour is acceptable and what kind of behaviour is not. Norms prescribe and proscribe behavior in specific circumstances and have long been considered to be at least partly responsible for regulating social behavior [10, 21]. Moreover, groups can have collective norms which are not written down but generally understood by each member. This type of norms is called implicit norms [4] and are of particular interest to our research, because on social networking platforms, generally all norms are implicit [15], despite various etiquette guides available online [20].
It is useful for our research to find out which kind of norms play the most important role on social networking services such as Facebook, because it is then possible to see how these norms determine what social media users think about others that violate privacy and share content without permission and whether they act upon it.
Research of Caitlin McLaughlin and Jessica Vitak learns us that norms on social media evolve as more and more groups of individuals make use of social networking services and as these users themselves move through life [15]. Within their research, they asked Facebook users to talk about the 'unspoken rules' that guided their Facebook use in order to get a picture of the implicit norms occurring. Norms identified this way were:

- Norms regarding friend requests;
- Norms regarding communication;
- Norms related to posting and tagging pictures;
- Norms regarding expectancies for friends' behaviour.

For our own research, we consider the last three types of norms as being relevant. First of all, it appears that norms of communication differ by whether the communication occurred through a public or a private channel. Sharing information with members within their network, like jokes, birthday wishes and other short messages should be done using features for public interaction, while in order to share personal information like a home address or organizing an event for a smaller group, private channels were more likely to be chosen. Additionally, a norm was present indicating that for longer messages to or entire conversations with an individual, private channels were the way to go.
When it comes to norms related to posting and tagging pictures, we learn that most pictures worthy of posting are pictures of people themselves and their friends at vacations or events. By doing this they attempt to create a positive impression towards other users, since users generally form a feel of someone by looking at ones' pictures.
Finally, the norms of expectancies for friends' behaviour on social media are that one is considerate towards his or her friends and friends are considerate back as well. This norm of consideration prevents friends posting inappropriate content to each other's pages.

Even though these implicit norms are generally being respected on a social networking service like Facebook or Twitter, there will always be users that do not adhere them and thus violate the implicit norms in place. In order to minimize the inappropriate behaviours that these norm violations entail, the concept of social control is often very helpful. We can define social control in the same way as research of Peggy Chekroun and Markus Brauer, consistent with past literature. They define social control as being "any verbal or nonverbal communication by which individuals show to another person that they disapprove of his or her deviant (counternormative) behavior" [5]. It has been shown that social control is present in 'offline' situations by many researchers already, like for example Emile Durkheim [9]. Moreover, its effectiveness can be seen, because when someone throws waste in a pond of a public park for example, in most Western societies this littering is considered to be wrong behaviour [12] and people that see it happening will often communicate disapproval in some way. This disapproval then encourages that others conform to the social norms. If we transfer this reasoning to the social web, (public) interventions for users that misbehave on social media platforms could potentially be a really strong weapon to

reduce the amount of misbehaviour as well and this way it could encourage good manners on the social web. However, we currently do not know to what extent people actually intervene others on social media when they disapprove of such deviant behavior (1.2) and we need to explore to what extent intervening is the norm.

2.2 Undesirable behaviour

Obviously, it is important to see how the norms identified on social media predominantly determine what kind of online behaviour is regarded as being unacceptable. Researchers Hooper and Kalidas [11] have done so and explored what behaviour is regarded acceptable and what is regarded unacceptable by youth and adolescents. This provides us with a list of these behaviours and other types of content that occurs among college students on social networking sites and in particular on Facebook. The list of unacceptable behaviours contains the following types:

- Rude or offensive postings;
- Embarrassing postings;
- Coarse language;
- Too much, too intimate and too detailed personal information;
- Unprofessional content;
- Breaching others' privacy; (content with sexual or illegal nature.)
- Randomly requesting friendship;
- Pestering a friends' friends;
- Stalking.

This shows that social media users breaching others' privacy, for example by sharing sexual or illegal content, is something other people really dislike to see. Therefore, it also displays once more why it is important to find out to what extent social media users try to intervene people performing such behaviour.

Hooper and Kalidas also found that users on social media often observe others' behaviour and then copy what they perceived as acceptable and unacceptable. Moreover, personal beliefs and values guides them in their ideas about whether something is acceptable or not.

2.3 Contextual privacy

Besides understanding what norms users apply on social media and what kind of behaviour is considered undesirable, it is also useful to understand what people's privacy expectations and privacy practices are and how these are determined. This can be better understood by making use of the theory and framework of contextual integrity. This theory designed by Helen Nissenbaum states that our beliefs about what is public and what is private are being affected by social expectations. Moreover does it state that such expectations vary with specific contexts [16]. It is not a full definition of privacy but it helps in understanding why certain patterns present in communication and information between people provoke public outcry in the name of privacy. Adapted to our own research this means that social expectations resulting from the social norms governing information flow, affect whether we think a certain type of information can be shared online with others without any problem or whether sharing this information should stay private and should not be exposed to others. This makes that understanding these norms (2.1) is a crucial starting point for understanding whether sharing a certain type of information is appropriate or not. As mentioned earlier, also the context in which the expectations users have about privacy appear plays an essential role in determining privacy practices. To give an example, usually we may be more comfortable sharing personal silly pictures with friends in a small WhatsApp group, but sometimes we may reveal surprisingly personal pictures to strangers we just met over Discord. This shows that the 'rules' which people follow for managing their privacy vary by situation.
Adding to this, from the body of privacy research we generally learn that norms and behaviors regarding whether something should stay private or can be told in public differs across cultures, but also between people of the same culture. Privacy concerns can even dramatically vary for the same individual over time [1]. However, if we consider privacy behaviour as context-dependent then this dilemma of what to share and what to keep private becomes something universal.

2.4 Types of reactions

Now that we know the different types of undesirable behaviour present on social media and explored the concept of contextual privacy, it is also important to explore different types of interventions possible when such unacceptable privacy breaching behaviour occurs. The interventions social media users apply when norms of inappropriate content and behaviors are violated have been explored by Yasmeen Rashidi, Apu Kapadia, Christena Nippert-Eng and Norman Makoto Su. Within their research they analysed

sanctions that allow individuals to shape community norms on appropriate privacy respecting behaviours [18]. From their research, we learn that these sanctions can be classified into different categories, namely on- and off-site sanctions (who performs a sanction), individual and collaborative sanctions (where do they sanction) and visible and invisible sanctions (how apparent is the sanction to the violator). The different sanctions that participants of the research indicated to be using for managing content sharing and privacy on social networking services are:

- Confronting;
- Gossiping and complaining;
- Passively reacting by ignoring/letting it go;
- Removing the violators/violated content;
- Implying disagreement by not actively approving;
- Uncoordinated and coordinated attacks using comments.

This list is useful, because it gives an idea about possible reactions and intervention methods users of social media could use. Therefore, we will be able to compare our results with those and see whether the participants of our research also make use of certain sanctions listed here.
We also learn that young adults are aware that imposing sanctions can be detrimental for those that want to avoid conflicts and maintain their relationships [18]. Therefore, they often choose for interventions which are indirect and invisible, even if they fail in expressing disapproval towards the violator.

2.5 Bystander effect

We have seen all these different types of sanctioning strategies that users on social media could possibly use. In certain situations these kinds of sanctions and interventions are being used by users on social media, but in other situations they are not. We have already seen how norm evaluation and contextual privacy can play a large role in deciding whether people ignore deviant behaviour or show disapproval. However, another possible effect which could play a determining role is the so-called 'bystander-effect'. This effect states that bystanders of a certain event or behaviour are less likely to intervene when there are more witnesses to an emergency [7, 14]. In later work we find that the bystander effect is not restricted to emergency situations only, but can occur in a variety of settings, like answering the door [17]. For our research this would mean that users of social media which come across unacceptable behaviour of other users are less likely to intervene, because of the fact that basically any user could potentially read it if they would intervene and because of a diffuse sense of responsibility.

The bystander effect has already been explored many times in the context of cyberbullying. For example, from the research of Leping You and Yu-Hao Lee we learn that when it comes to intervening if someone is being cyberbullied, users' intentions to intervene increase with the bystander numbers, but drop after a ciritcal point [27]. Generally, this research concludes that the relationship between bystander numbers and intentions to intervene depends on factors such as anonymity and the type of interventions. Anonymity plays a strategic and a cognitive role determining intentions to intervene in an online bullying context.
Research paper DiFranzo et al. explored whether the interface design of a social media platform influenced bystander intervention when it comes to cyberbullying [8]. This research indicates that bystanders are more likely to intervene indirectly than directly, which we have also seen in chapter 2.4. Moreover, it shows us that information about the audience size and viewership is able to increase the likelihood of reporting cyberbullying posts through serial mediation of public surveillance, accountability and personal responsibility. Therefore, as a design implication it suggests to display audience size and public surveillance cues, such that cyberbystanders accept responsibility for intervention.
In conclusion, we can clearly state that the bystander effect could potentially also be an influencing factor when it comes to whether users of social media intervene other users, violating people's privacy by sharing content of them without having permission for doing this.

2.6 Expectancy violation theory

When someone's behaviour deviates from what was expected by others, you could say that the expectations of this behaviour are violated. When someone's expectations are violated, this violation is judged as being either a positive or a negative violation. Moreover does this deviation influence the so called 'communication outcomes', which basically means it influences the way you respond. Typical examples of communication outcomes are comprehension, attitude change and attraction. All of this is basically stated by the Expectancy Violations Theory of Judee K. Burgoon [3]. As explained earlier already, social media users all have their own norms and values which they apply online. It can occur that these norms are being violated by other users of social networking services. According to the research of McLaughlin and Vitak mentioned earlier, if this happens then the closeness of the relationship you have with the norm violator plays an essential role in determining how you would react or sanction this violator [15]. Simply put, this means that when friends of you violate your norms on social media, you are more likely to confront them than when acquaintances violate your norms. Moreover, it turned out that for acquaintances the severity of the violation primarily determines how you would react.

2.7 Second order cybernetics

A theory or science which is concerned with structures, constraints and possibilities of purposive and regulatory systems is known as cybernetics [26]. Wiener defined it as "the science of control and communication, in the animal and the machine" [24]. Cybernetics basically sees a system or algorithm, the users of this system and the designers of it, as one unit. As a concrete example you could see a user of Facebook, Facebook itself offering its services to this user, and Facebook as developer of new features based upon the behaviour of this user, as one system. Within this system, both the user and the algorithm have a responsibility to make the interaction between the different parties as useful and functional as possible. Cybernetics treats ways in which people behave [19].
A meta-field is called second-order cybernetics or design cybernetics [13]. Generally explained, this field is about cyberneticians observing these systems and seeing whether these interactions actually happen and what would eventually need to change to make events and interactions run smoother. It provides us with feedback and can learn us things about the design of a system and how this is able to steer human behaviour. Because of this, second order cybernetics may produce design implications for systems, also on the social web. Simply rendered, for our research this means that the design of the interface of a social media platform has an effect on how users of that platform interact and make use of it. On the other hand, their behaviour also influences the social networking services, which from time to time add and remove different kinds of functionality.

Chapter 3

Research

Now, in this chapter we will discuss the research that we carried out in order to obtain more insight in answers to the research and sub question, as stated in section 1.2
First we will discuss the methodology we applied while carrying out our research and explain our research strategy and the data generation methods we used. We will also discuss the type of data analyses used and specify how the quality of our research results will be ensured. Next, we will describe our results and present them in a clear and cohesive way. Finally, we will also provide a discussion which will give meaning to the results we found and put them into the context of the theoretical framework we discussed before. Moreover, we will highlight potential limitations that our research entails.

3.1 Methods

In order to find answers to our research question and sub-question, we conducted in-depth interviews with 15 different people. The interviews were semi-structured and were held via the online videotelephony and chat services Zoom and Discord. Additionally, the audio of the interviews was recorded such that processing and analysis of the generated data was possible. All participants had been informed of this in advance and gave their consent. The main reason that we chose to work via these online services and did not interview people physically at a certain location is because of the currently ongoing global COVID-19 pandemic. Moreover, this is a convenient way to contact and interview people that live across the country. Each interview was conducted individually in a one on one conversation and participents were recruited via our own student network and the networks of friends and acquaintances. There was one eligibility criteria for participation, which was that participants must be frequent or active users of currently popular social media platforms, such as Facebook, Twitter or

WhatsApp, which can also be considered 'online public places', in contrast to small groups of friends or private communities.

The interviews lasted between 19 and 28 minutes ($M = 23$ min, $SD = 3.18$ min) and out of the 15 participants, 6 identified as male, 8 identified as female and 1 did not identify itself with any gender. Most of the participants (9) were graduate students from different fields of study and were aged between 18 and 25 years old. 3 participants were undergraduate students and aged between 16-17. The other 3 participants were older than 50. The majority of participants is student as this group is generally more active on social media and was more approachable to us while doing the research. Out of the participants, 12 reported they had been using both Twitter and WhatsApp. Most of them (8) were also using Instagram and Facebook. Finally, some (8) also reported they had been using Snapchat and (7) TikTok. Every participant reported making use of social media on a daily basis, where most of them (11) even reported using it a few times per hour.
The reason we made use of in-depth interviews, is because our research is an exploratory research and conducting in-depth interviews is a qualitative data generation method in which you are better able to learn about people's reasoning and intentions, especially when it comes to how they behave on the social net.

Before we actually started conducting our interviews, we first composed a framework of questions and themes that we wanted to explore during the interviews. This way we made sure that everything we wanted to explore was actually asked and the interviews did not deviate too much from the main topic and research questions. This framework was then tried out in a few 'practice interviews' in order to get some experience interviewing and to find out about possible problems with the framework, like whether it was missing data which we would actually like to know and whether the order of the questions in the framework made sense. In the end, these 'practice interviews' did lead to some updates of the semi-structured interview protocol.
Now, within this pre-prepared interview structure (see appendix A), next to some other questions, we basically created two cases. The first case asked whether the participant could remember one or multiple events from the past in which online friends breached other people's privacy by sharing content of others without permission (also see chapter 1.2). This question was then elaborated more on by using concrete examples to make sure participants understood the situation properly. Once they recalled such event, they were asked explicitly to answer the next questions while having that specific event in mind and explaining what they did in that situation. From time to time, we asked for details and asked other follow-up questions, such

that participants helped us to understand what occurred from their perspectives. This way, we made use of the critical incident technique [6] and made sure people reported what they actually did in a specific situation, instead of reporting what they would want to do in a certain situation or what they believe would actually be the best act to perform. This of course gives more accurate data about one's behaviour.
If the first case did not apply and the participant either did not experience such event before or could not remember it anymore, the second case applied. Here, participants were asked to just imagine a situation in which online friends breached other people's privacy by sharing content of others without permission. In order to try and minimize the effect of the aforementioned problem where participants would report acts they think would be the best to perform in such situation here as well, we asked to take a specific online friend and a platform in mind while answering our follow-up questions. Depending on our questions, we also switched what kind of online friend the participant should try to think about.

After we were done conducting the interviews, we rewatched the recordings and transcribed every one of them in a very detailed manner, using word processor Microsoft Word. Finally, in order to efficiently analyze the different answers that participants had given, we made use of qualitative data analysis & research software ATLAS.ti 8, and we created 34 different codes and 279 quotations in order to categorize and structure the various results. During the coding process, we chose to make use of 6 general overarching themes and each time we came across a participant that described something we did not encounter before, we created a new code and identified under which theme this code would fit best. This way we made sure we did not miss any data the respondents provided and were able to structure our large data set and comparable information such that it became overseeable.

3.2 Results

Now we will present the results that we derived from the transcriptions of the in-depth interviews we conducted. We will do so on the basis of our codes. Sometimes, we will make use of quotations from the interviews that support our results and which increase clarity. Please be aware of the fact that these quotations were originally spoken in Dutch, but are now translated to English in order to improve the flow of this research paper.

The overarching themes that we used within our coding process were set up as code groups and named 'Demographics', 'Reaction types', 'Reaction reasons', 'Relationship with violator', 'Influencing factors' and 'Bystander effect'. Moreover, we also came up with a few stand-alone codes, like 'Opin-

ion about behaviour' and one about the reporting mechanism'. We also identified whether participants' answers were based on a past experience or based on what they would do in a certain hypothetical situation that was brought up. This structure can be seen in figure 3.1:

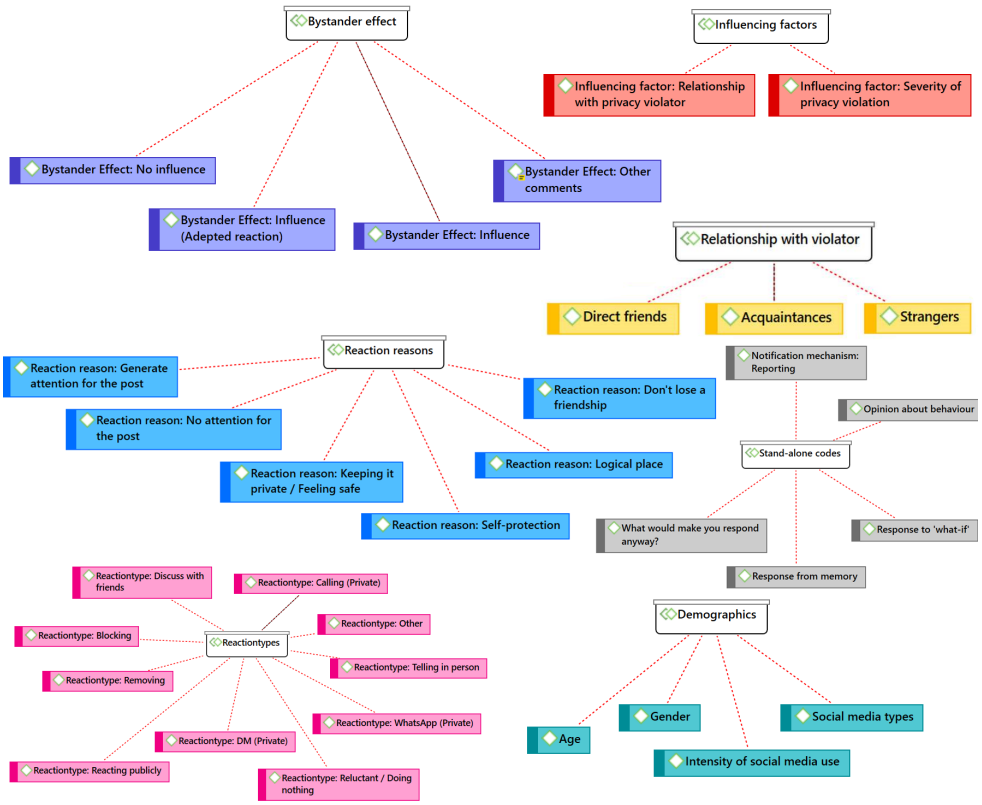

Figure 3.1: *Code structure*

In section 3.2.1, we will discuss what the participants of our research their opinion was about users on social media violating the privacy of others by sharing content of them without having the permission to do so. Then, in section 3.2.2, we will present different ways of reacting and intervening that participants indicated during the interviews. Subsequently, two important factors that determine and largely influence whether and why participants react and intervene in a certain way will be presented. In section 3.2.3 we discuss how the relationship one has with the privacy violator is of influence and in section 3.2.4 we discuss how the severity of a privacy violation is of influence. Next, in section 3.2.5 we show whether the reporting mechanism which is often present on social media was being used by participants of our research. Finally, in section 3.2.6 we describe what respondents said about the bystander effect and the influence they believed it has on them.
In figure B.1, a concise overview of the statistics of the results can be found.

3.2.1 Opinion about privacy breaching behaviour

Before we asked participants how they reacted or would react, we investigated what participants generally thought of online friends and social media users violating the privacy of others by sharing content or information without having their permission. The reason we did this is because people could potentially have no problem with such behaviour on social media and therefore react in a certain way. However, this turned out to be not the case. Out of the 15 participants, 14 said that this kind of behaviour is unacceptable and inappropriate. Some said it made them angry while others said they found it almost pathetic. In one of the interviews we simply forgot to ask this question, which is why we also do not have an answer of one of the participants.

"Well, that doesn't make any sense. I really think it is worthless. I really can't stand it when people behave like that. Definitely unacceptable."

Also an interesting comment of one of the participants was that for this person, it really depends on the severity of the privacy violation, whether this kind of behaviour is considered unacceptable or a type of humor that is right on the edge. We will discuss this difference later in section 3.2.4.

In conclusion, basically all of our participants find privacy violating behaviour unacceptable and the decision to not intervene does not relate to users approving the behaviour.

3.2.2 Reaction types

After we asked for participants' opinion on this privacy breaching behaviour, we asked them questions about how they would react and whether they would intervene in certain situations. The answers participants gave were based on past experiences or imaginary 'what-if'-scenarios. We will now describe the different types of reactions and intervention methods participants indicated.

3.2.2.1 Privately

Most of the participants, 12 out of 15, indicated that they would most likely react in private or reacted in private during a past experience. Different types of reacting in a private way have been mentioned. Which one someone goes for depends on whether the phone number of the person violating the privacy of others is known or not. Please note that one participant may have indicated using multiple types of reacting privately.

- Sending a direct message (DM) via social media. (12 participants)
- Sending a message via WhatsApp. (3 participants)
- Calling. (2 participants)

"That depends on whether I have his or her phone number. Otherwise I would send a DM."

Reasoning
Several reasons for intervening people that breach others' privacy have been indicated. Many participants explained that they did not want to generate more attention for the post violating the privacy of others. If they would reply publicly, the post would get more interaction and would get pushed more by the social networking site to for example followers that follow the intervenor. Moreover could it be possible that others will find this post via your profile and this way you participate in sharing the content without permission as well.
One participant said that they reacted in private if they knew the violator very well. This way asking why he or she had posted such content without permission would become easier, as the violator would feel more safe, compared to publicly giving a honest reply.
Another participant said that he considered the direct messages section simply as a more convenient and more logical place to reply, due to the fact that having conversations there runs smoother.

3.2.2.2 Reluctant

Many participants, 13 out of 15, also indicated that in certain situations they would either not reply at all, or be reluctant with their reply. Often they then choose to make use of the reporting functionality that many social networking sites have implemented. We will discuss this later in section 3.2.5 as well.

Reasoning
One reason to be reluctant with intervening that was mentioned often is self-protection online. Respondents explained that they were afraid of becoming a target of the person breaching other people's privacy as well. Also, they were scared that other social media users would think they regularly engage

in negative posts and that this would affect their image in a negative way on social media.

"If I had responded to it online, it would have caused more trouble for me than for what it was actually about."

Therefore, if you would be able to reply to someone anonymously - especially to strangers - intervening publicly might become easier for social media users. This was also mentioned by one of the participants:

"If you're anonymous or something. Maybe this sounds silly, but if I were to intervene myself then it feels very personal as well. So if I could just send that person a message anonymously, I think I would do so faster."

Another reason to be reluctant respondents mentioned was to minimize the amount of interaction and attention the post would get, as explained already in section 3.2.2.1 as well.

3.2.2.3 Publicly

There were also a few participants, 4 out of 15, who explained that they would react publicly or did do so in the past already, for example by commenting under the post itself that breached someone's privacy.

Reasoning
The main reason respondents gave to react publicly on social media, is that they considered violation of privacy and sharing content without permission as being such unacceptable behaviour, that it is important that as much attention as possible is generated for it. Other people should see it too and ultimately, should intervene as well. This way the privacy violator's image will be affected and the chance of one doing it again should decrease.
Another reason that was mentioned by a participant is that once a discussion started somewhere and people think something about a certain post, it is important that the entire conversation is held underneath that post such that others also see the intervention happening and do not suddenly miss parts of the interventions. This way other social media users will finish up the story themselves.

"I am always like okay, at that place the discussion started, then we will finish it there as well. (...) I don't like it if people continue their reaction 1 on 1, because (...) everyone has been able to read the beginning. We should finish it there before people finish the story themselves, because then it will become a mess even more."

3.2.2.4 Blocking or removing

Next to reacting with a message or a conversation, whether this is done privately or public, some participants also carried out certain actions in order to intervene. The first one indicated by 5 out of the 15 respondents is blocking or removing someone from social media. If you follow someone, which suddenly starts posting for example nude photo's of someone and starts violating other people's privacy, simply removing or blocking this person could be an effective approach. This way you increase the chance that this post will be seen by a minimal amount of people and it also prevents a discussion from happening.

3.2.2.5 Discuss with friends

4 out of the 15 participants indicated that they also sometimes feel the need to discuss it with friends. They then tell their friends about what happened and ask them to report the privacy violating post as well or also intervene the person violating the privacy of others.

"Furthermore, I sometimes also discuss it with friends of the person violating others' privacy or with friends of myself, if they also know that person."

3.2.2.6 Other reactions

When analyzing the transcriptions of the conducted in-depth interviews, we also found a few actions that were only mentioned once overall. These other types of reactions are listed here:

- Placing multiple posts at once with an opposite purpose, if the goal of the initial post is to spread hate or make fun of someone.
- Calling the police, if a privacy breach can be incredibly impactful.
- Telling your parents, if the person violating the privacy of others is an acquaintance. Chances are then that your parents know the parents of the violator and once they tell them what happened it is possible that they will intervene the violator.

To summarize this section, participants indicated different types of reactions and intervention methods they used on social media. Each reaction type has a reasoning of why participants use it and the different types mentioned can be classified into 6 different 'overarching types':

- Intervene privately;
- React reluctant or do nothing at all;
- Intervene publicly;
- Block or remove;
- Discuss with friends;
- Other types.

Out of these different types, private intervention as well as reacting reluctant was often a preferred way of acting for our participants.

Overview
In the following diagram, all different reaction types and the amount of times respondents indicated them are shown:

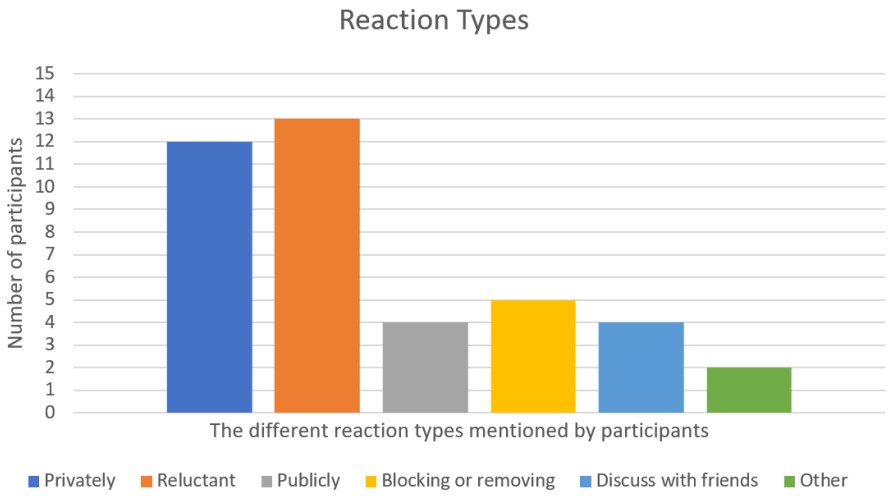

Figure 3.2: *Reaction types*

3.2.3 Relationship with the privacy violator

One important factor that determined why participants decided to react in a certain way and which emerged from the conducted interviews is the strength of the relationship one has with the person exhibiting the privacy violation. We will now describe how these different strengths of bonds influence the way respondents react.

3.2.3.1 Direct friends

First off, 12 participants of the conducted interviews described that it would be easier for them to intervene someone sharing content of someone else without having his or her permission, when this person is a direct friend. They would do this in private then.

"If it is a close friend then I would take action quicker."

"Probably privately, um... So you don't make fun of someone you know in front of a whole group, so to speak."

Reasons mentioned for this are that you know direct friends well enough to make an estimate of how they will reply to you. You are better able to listen to each other and you know each other's limits. Also, you know how to talk to each other.

"Because I think someone I talk to regularly, I would be more likely to address and someone I almost never talk to less likely to address."

However, on the other hand there where also 3 participants who said being direct friends with someone complicates intervention. When you are good friends with someone, you rather do not want to have major discussions or risk losing the friendship you have. You want the bond to remain good. Therefore, you choose to react reluctant.

"I did not really actively go against it, because I also did not feel like losing that friendship or something."

This shows that even though the relationship someone has with the privacy violator influences how one decides to react, the actual way in which one reacts in the end is quite personal and can differ from person to person.

Overview
In the following diagram, the different indications participants gave are shown, together with how many participants described them:

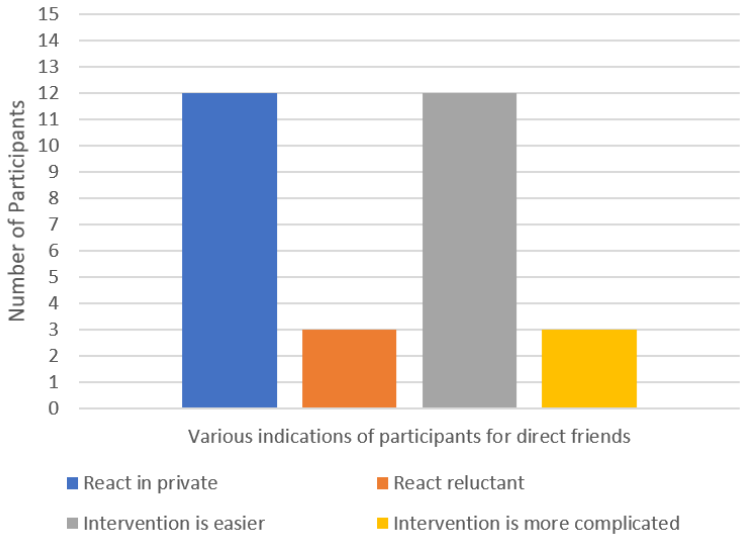

Figure 3.3: *Indications participants gave if the privacy violator is a direct friend*

3.2.3.2 Acquaintances

When it comes to acquaintances, not much seems to change compared to being good friends with the violator. Also here, participants indicated that they would most likely send a direct message in a private chat in order to intervene.

"Yes, if I had known the person or spoke to them occasionally I would have said something about it, because then you kind of know what that person is like."

One participant explained that he or she would address someone less quick if this person breaching someone's privacy would be an acquaintance, compared to how fast he or she would address a direct friend.

When talking about the reporting functionality which many social networking sites contain, 2 participants indicated that they were less likely to report acquaintances or would not report acquaintances at all, if they breach someone's privacy. They would report strangers however.

"I do believe reporting is effective, but then I would not report acquaintances and I would report strangers."

Overview
In the following diagram, the different indications participants gave are shown, together with how many participants described them:

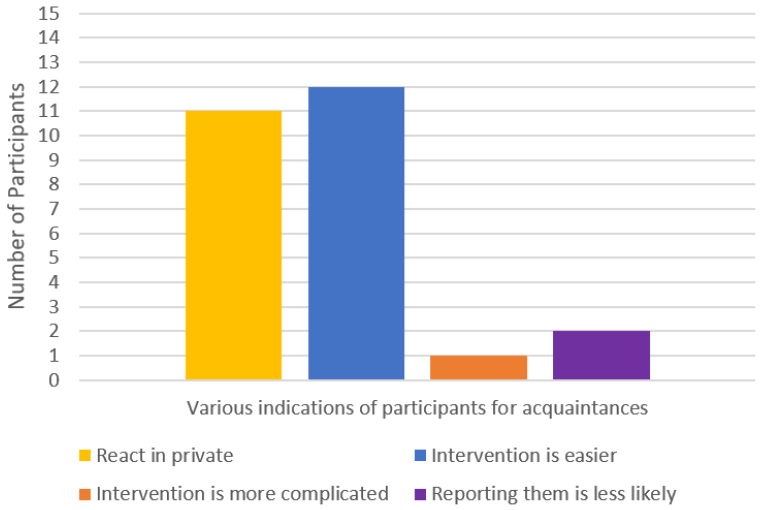

Figure 3.4: *Indications participants gave if the privacy violator is an acquaintance*

3.2.3.3 Strangers

When it comes to strangers, participants mainly indicated two different types of reactions they would give to people violating the privacy of others on social media. First, 11 participants said they would react reluctant or not react at all. The reason for this is that if you do not know the violator personally and you also do not know the person who's privacy is breached, it is difficult to assess whether the violator crossed the boundary or not and actually breached someone's privacy. Because of that determining whether intervention would be needed becomes more difficult and it is then easier to be reluctant with your reply.

> *"Look, if I do not know the person, I am not the one to address him. I would still try, but it is not really... Why would he listen to me, you know?"*

A different type of reaction which 4 participants mentioned is commenting and reacting publicly. An example reason to choose for this approach is that they do not know the violator personally and reacting publicly is a smaller step than reacting in a private message or call.

"Because I do not know the person, therefore the personal approach... Yes, I simply never DM with that person, so it doesn't really matter where I respond."

Overview
In the following diagram, the different indications participants gave are shown, together with how many participants described them:

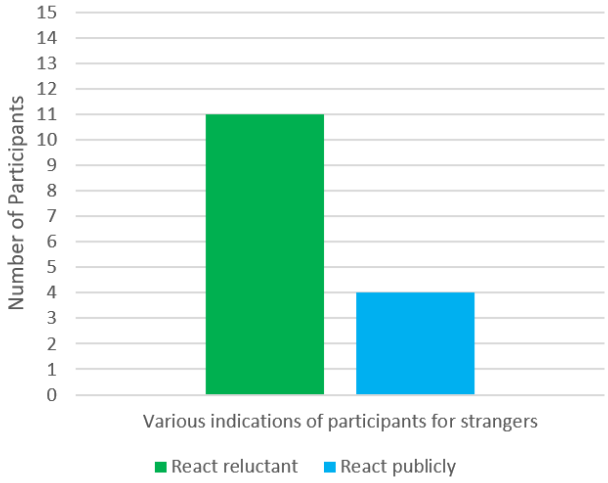

Figure 3.5: *Indications participants gave if the privacy violator is a stranger*

Summarized, we learn that most participants indicated that intervening direct friends or acquaintances is easier to do than intervening strangers. Additionally, they would react in private most of the time, unless the privacy violation was performed by a stranger. In this case, reacting reluctantly is the reaction type which is chosen most often. Some participants however would actually reply publicly in such case. Also, the minority of respondents indicated that intervention could actually become more complicated when it considers direct friends.

3.2.4 Severity of the privacy violation

Another factor that emerged from the conducted interviews is the severity of a privacy violation. 11 respondents explained how their reaction on online friends and social media users violating the privacy of others by sharing content or information without having their permission, was also depended on how bad a violation was. For example, there is a difference between someone sharing your date of birth online, while you actually tried to keep this secret and someone sharing a video of you doing stupid things while being drunk. In the first case, it is not very likely that people would feel the need to intervene, while in the second case there is a much higher chance they would react, even if it would be only by making use of the report functionality. Additionally, participants said that their reaction was also "depended on the situation" in which the violation occurred.

"But I also think that it matters what actually the photo or post is. Like, if I suddenly see a picture of someone's face I would pounce on it a little bit less than if it is some nude picture, because I feel like that can still bother you later, more than a picture of your face (...). So, how severe or how big the impact is of the privacy violating post, I think that matters a lot."

Participants also mentioned that it is difficult to make your own judgments about when something is okay for a certain person and when a certain post is violating someones' privacy. This boundary is different per person and sometimes you might make a wrong judgement. One participant explained they often made these judgements by imaginatively placing themselves in such situation and determining whether they would accept a certain post.

"Yeah I think that, for me it is like a personal decision or something. So I would simply place myself in the victim's shoes all the time and then judge whether I would accept it if someone else did that to me or not."

In conclusion, the severity of a privacy violation was an important factor influencing whether respondents chose to intervene or not and to what extent.

3.2.5 Reporting mechanism

Many, if not all, social networking sites and platforms today include a reporting functionality. This way this site or platform can be notified by its users that something or someone is violating the rules of the social networking site or platform. Action can then be taken by them as well. During the interviews, we also asked the participants whether they had ever used a reporting mechanism in the situation of someone breaching the privacy of other people without their permission. Interestingly, it appeared that only 2 participant had not used this functionality from time to time in such situations.

"But I did report that account several times, because of the privacy violation, etcetera."

This clearly shows that the reporting functionality present on social media already is being used very often and is a useful feature in the design of different platforms.

3.2.6 Bystander effect

During the in-depth interviews, we also asked the participants about whether they thought the bystander effect influenced how they would react and whether they would react at all on social media, when they encounter privacy violation. Of course, we first explained to them what this bystander effect actually is. Out of the 15 participants, 12 said the bystander effect does not really influence them.

"That hasn't played out for me, because I really don't care one bit about what other people think or do. I'm completely doing my own thing, especially with things like this as well."

However, out of these 12, 8 said that they do take into account the fact that others can read along and can see what you are posting when formulating a message meant for intervention. You are not alone on the social web.

"There may be thousands of people looking at that post, suppose we take the Twitter example and an inappropriate photo is posted. I post my opinion about it, but I know that this opinion can be seen by others. (...) Then I'm going to explain in a normal way that this is just very inappropriate, because if I behave aggressively then people from other societies or friends of me who read that message might develop bad thoughts about me. (...) So I actually make sure that I first think about it logically before I post a message."

There were also 3 participants that indicated the bystander effect does actually influence whether they do or do not intervene on social media. They indeed said that if no one said anything, they also kept themselves in the background.
Next to this, one participant explained that he or she would have replied if other social media users would have said something about it as well.

Overview
In the following diagram, the different indications participants gave are shown, together with the amount of participants that described them:

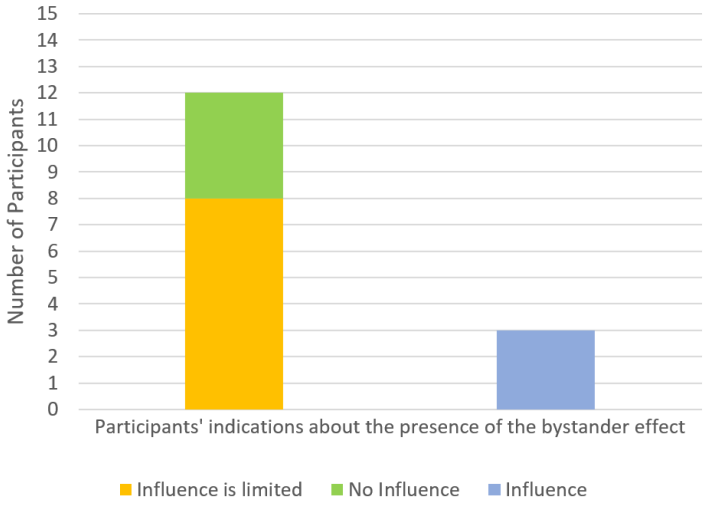

Figure 3.6: *Indications participants gave about the presence of the bystander effect*

3.3 Discussion

Now it is time to discuss and analyze what the results we have just presented actually imply. In subsection 3.3.1 we will first interpret the results that we found and we will also relate them to the related work, as presented in chapter 2. Next, we will come with a general concept of possible functionality as design implication for social networking platforms in subsection 3.3.2. Finally, we will end this section with a description of the limitations that have to be considered with our research in subsection 3.3.3.

3.3.1 Findings

Since our research consists of 2 parts; our main research question and a sub-question, we will discuss our findings here in 2 parts as well. In section 3.3.1.1 we will discuss our findings related to the main research question and in section 3.3.1.2 we will discuss our findings related to the sub-question, which both are defined in chapter 1.2.

3.3.1.1 Intervening on social media

3.3.1.1.1 Types of interventions

Within our investigation, we have found that there are different actions and different ways in which people react towards their online 'friends' on social media. The most used types of reactions people give when they see a different user on social media breaching the privacy of someone, are either contacting the violator privately or reacting reluctant and basically ignoring that it happens. When one choose to reply in private, in most cases they choose to send a direct message over the social media platform where the violation took place. Sometimes, they also choose to send a message via WhatsApp or call the violator.
Another type of reaction which we have found is reacting publicly. However, due to the minor amount of participants indicating they make use of this type, we conclude that this is not a preferred way of reacting among social media users.
Next to this, other types of reactions and actions that we have seen and which users of social media use are blocking or removing the violator and discussing what happened with other friends. Additionally, other actions like calling the police or telling what happened to your parents have been indicated.

Many of these different types of reactions and intervention methods match the different sanctions we have discussed earlier already in section 2.4 of the related work. First off, we can conclude that indeed many users of social media choose to confront someone violating other people's privacy, as the

research of Rashidi, Kapadia, Nippert-Eng and Makoto Su also indicated. This research did mention that such confrontation in most cases takes place 'off-site', but considers private messages as being off-site as well [18]. Therefore, this finding is consistent with our results presented in section 3.2.2.1.
Next to this, we have also seen reaction type 'gossiping and complaining' a few times within our results. With this type of reaction, people want to avoid an online argument with the violator and therefore choose to discuss what happened with people that share their point of view about posts with which they disagreed [18]. This corresponds with the reaction type mentioned by our participants of discussing what happened with other friends as can be seen in section 3.2.2.5.
Finally, also passively reacting by ignoring/letting it go, removing the violators/violated content and uncoordinated and coordinated attacks using comments, found within the same research paper and discussed before in section 2.4, have been shown present within our results. Passively reacting by ignoring/letting it go matches the reaction type of reluctance described within our results in section 3.2.2.2. Removing the violators/violated content corresponds to our participants indicating in certain cases they remove or block privacy violators on social media (see section 3.2.2.4). Lastly, the uncoordinated and coordinated attacks using comments with which online users give visibly direct feedback to inappropriate posts is consistent with our result in section 3.2.2.3 which learns that people react publicly in certain situations.
The fact that the identified types of interventions people use on the social web we identified correspond with the results mentioned earlier within our related work, increases our confidence in the results that we found and improves their representativeness.

3.3.1.1.2 Influence relationship privacy violator

Our present findings also suggest that participants decide in which way they intervene a person breaching others' privacy by evaluating how well they know this violator and what their bond is. The majority of participants indicated that if the violator is a direct friend, they would actually take action and intervene. However, they would do so in a private way (3.2.3.1). In the case that the violator is an acquaintance, most participants said they would actually intervene as well here using the same approach (3.2.3.2). When it comes to strangers however, most participants said they would react reluctant or not react at all (3.2.3.3). Only a few participants said that intervening strangers is easier than intervening direct friends, because of the potential risk of affecting your friendship which does not exist when intervening people you do not know. However, we consider the amount of 4 respondents indicating this as being not major enough to draw strong and reliable conclusions. Therefore, we conclude that generally, people take

action and intervene violators of privacy quicker when they are direct friends or acquaintances, than when they do not know the violator.
These findings are in line with the expectancy violation theory, which we have discussed already in chapter 2. According to this theory, when one's norms - or expectations - are violated, this deviation influences the way one responds. In particular does this theory mention that closeness of the relationship with a norm violator plays an essential role in the kind of reaction one gives, as we discussed in section 2.6. Following the expectancy violation theory we learn that if friends violate your norms on social media, you are more likely to confront them than when acquaintances violate your norms. As can be seen, this is also what we see reflected in and what we concluded from our own findings. However, one difference is that according to the expectancy violation theory acquaintances are less likely to be confronted, while this is not apparent from our results.

In conclusion it turns out that users of social media do actually intervene their online 'friends' most of the time. However, when they do so they often do this in private. This is why interventions are not found on social media as much as the amount of times it actually happens. People do actually intervene their online 'friends'. When it comes to strangers however, intervention is less common. This is in line with the present literature known as well.

3.3.1.1.3 Influence severity of privacy violation

From our findings, we learn that not only the closeness of the relationship with a norm violator is an important factor for social media users when deciding whether they should intervene or not. Also the severity of a privacy violation is influential. When someone shares your date of birth online, this is a different kind of privacy violation than when someone shares an embarrassing video of you (3.2.4). Additionally, our results within this section show that it is often difficult to judge whether a certain privacy violation is considered a privacy violation by the person who's privacy is being violated as well.
All of this is in line with the concept of contextual privacy, discussed in section 2.3. This concept shows us that both our beliefs about what is public and what is private and also the context in which expectations people have about privacy appear, influence whether someone considers sharing a certain type of information as being appropriate or not. Therefore, if something is shared by someone else which does not match our own social expectations, we would consider this as something that should have stayed private. However, since for these judgements we make use of our own beliefs, estimating whether certain behaviour is a privacy violation for someone else you do not know is rather difficult.

Next to our own beliefs, also the context influences our social expectations which we are also able to find back in our results.
The research of McLaughlin and Vitak [15] that we discussed in section 2.6 also mentioned that, in the case of acquaintances, the severity of the violation primarily determines how someone would react.

In conclusion we can state that the severity of a certain privacy violation is indeed an important influential factor when an individual is to decide whether and to what extent he or she should react on privacy violating behaviour on social media. Additionally, this matches what we already know from the literature and provides a reason to why intervening does not always occur on the social web.

3.3.1.2 Bystander effect

When exploring our sub-question defined in chapter 1.2, we need to look into our results about the bystander effect in section 3.2.6. These findings indicate that the presence of the bystander effects seems to be quite limited. When our participants where asked about it, the majority explained how they did not really take into account the amount of 'online bystanders' and that this effect did not influence them. They considered people breaching other people's privacy as being such a wrongful act that intervention should not be affected by the fact that others can read along. On the other hand however, participants did say that they take into account the fact that other people are able to read what they post online. Therefore, they often reacted more neatly than what they had in their minds, which shows that it is not entirely true that the idea that there are many 'online bystanders' present does not influence their behaviour and willingness to intervene at all. Consequently, we argue that the bystander effect was actually present, but only in a different and minimal way.
Looking at the related work in chapter 2, we see that in the context of cyber bullying, the bystander effect does play quite a role. Within our research we did not find any leads to intervention intentions increasing up to a certain point as bystander numbers increase, after which they dropped, which was concluded within the research of Leping You and Yu-Hao Lee [27]. Additionally, we do not see a reason to display audience size and public surveillance cues in order to make bystanders accept responsibility for intervention, as DiFranzo et al. suggests as a design implication [8]. This makes our findings about the bystander effect somewhat deviating from other literature. As our research was exploratory, it would be good if future investigation on the bystander effect in the context of social media and behavioural norms will be executed using a more in-depth approach.

3.3.2 Design implication

Now that we have discussed our results presented in chapter 3.2 and we related our findings to the related work written in chapter 2, we will provide a very general concept of a design implication, which could potentially increase to what extent people intervene their online 'friends' on social media, behaving undesirable by breaching other people's privacy and sharing content of others without permission (1.2). Increasing people's motivation and willingness to intervene others on social media is important, as it has potential to to decrease the amount of unacceptable behaviour that occurs on social media and encourages good manners on the social web.
Because of the fact that intervening among our participants happened the least amount of times when the person violating others' privacy was a stranger, it would be useful to make intervention in these cases more approachable. We suggest doing this by using anonymity as key principle. From our results presented in section 3.2.2.2, we learn that one of the core reasons people reply in a reluctant way is that they are afraid of becoming a target themselves as well of the person violating others' privacy. One participant even indicated that if they would have the possibility to reply anonymously, intervening could become much more approachable and easier to do. Moreover, from the literature on second order cybernetics - also known as design cybernetics - that we discussed in section 2.7, we know that the design of the interface of a social media platform has an effect on how users of that platform interact and use it. Therefore, implementing a new functionality can possibly increase the amount of people intervening violators of privacy.
A possible way of implementing such new functionality could be for example by adding an extra like or dislike button to each post on social media. The idea of this button would then be that it becomes possible for users to privately and anonymously signal someone when they think the one that posted the content should take an extra look to what they placed online. This way the person that posted the susceptible information can easily and anonymously be intervened in a private way by others. The threshold for using this new mechanism would be lower and therefore intervening others online would become easier. Obviously, this concept has to be researched more in depth in the future and whether it actually works still has to be shown. However, reasons why it has potential can already be given.
Firstly, this concept can be compared to functionality that is often present already on different web fora and is used to report a certain message to a moderator. In our case however, the report would not be received by an external party or the social networking service itself, but it would be received by the person posting the privacy breaching content. Also, we can expect that many users would actually make use of the new functionality, because judging from our results presented in section 3.2.5 we know that almost all

of our participants have used and still use the reporting feature that is often present on social networking sites already. This type of reporting signals inappropriate posts to the service itself which then participates, judges and sometimes intervenes as being an external party. In conclusion, adding such a new feature could encourage social media users to intervene in even more cases, which eventually increases social control (2.1) and good manners on the social web.

3.3.3 Limitations

We urge caution on prematurely generalizing from our findings, because of the fact that our participants were mainly graduate and undergraduate students under 26 years old. Nevertheless, we did find these young adult participants remarkably careful and adept in dealing with privacy violations by others on social media. Because of the fact that all our participants where intensive, experienced and savvy users of social media, we do believe that the results could potentially transfer to all social media users, regardless of age or other demographic category.
Additionally, because our interviews did not directly assume a specific social media platform, the results could also be appropriate for the range of different social media platforms. However, the fact that Twitter was often named as an example in order to concretize the different situations that people were presented with could mean that the results show a bias towards this social media platform.
Finally, it is important to keep in mind that participants - if they could not recall a clear moment from the past - were asked to describe what they would do in certain situations in relation to privacy. Because of the fact that people have a tendency to describe what they think they would they do or what they think would be best to do it is very well possible that if such situation really occurs, they react differently than the way they said they would react. This can be linked to the so-called privacy paradox [2], which states that users tend towards privacy-compromising behavior online, resulting in a separation between privacy attitudes and actual behaviour.

Chapter 4

Conclusions

Our research has been an exploratory research. It explores different ways in which interventions take place on social networking services and it provides us with important factors that influence whether intervention of other users violating the privacy of online 'friends' takes place or not. Moreover, it partly examines existing theories that have already been established in the body of privacy and intervention research and literature. Also, it provides a very general concept of design for new possible functionality applicable on social networking services, of which still has to be proven that it works and has to be worked out more in depth. Finally, our research also investigated the bystander effect and whether it is present on social media services when a privacy breach by a social media user occurs.
We will now present the most important conclusions that can be drawn from the preceding chapters:

- Firstly, our research demonstrates that intervening on social media is something which is very personal. Every individual has their own implicit norms and social expectations which they use to judge about privacy practices and violations. Additionally, we learned how severity of a privacy violation and the relationship one has with a privacy violator are of big influence in determining intervention.
- Secondly, we learned that people on social media do actually intervene their online 'friends' quite regularly. However, often you do not see that this happens, because most of the time the confrontation or intervention takes place in a private chat or conversation.
- Finally, it turns out that presence of the bystander effect in the form known within other literature is quite limited, even though it has been proven present in the context of cyber bullying. On the other hand, users of social media are actually being influenced by the presence of 'online bystanders', because our results indicate that people do change the way in which they act, react and interact on social media to a neater way, when intervening publicly.

We are optimistic about to the future of behaviour which is present on social media and are pleased that this research contributes to understanding mutual privacy and good manners on the social web!

Bibliography

[1] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Privacy and human behavior in the age of information. *Science*, 347(6221):509–514, 2015.

[2] Susan B. Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), Sep. 2006. https://doi.org/10.5210/fm.v11i9.1394.

[3] Judee K. Burgoon. A Communication Model of Personal Space Violations: Explication and an Initial Test. *Human Communication Research*, 4(2):129–142, 03 2006. https://doi.org/10.1111/j.1468-2958.1978.tb00603.x.

[4] Gary Burnett and Laurie Bonnici. Beyond the faq: Explicit and implicit norms in usenet newsgroups. *Library & Information Science Research*, 25(3):333–351, 2003. https://doi.org/10.1016/S0740-8188(03)00033-1.

[5] Peggy Chekroun and Markus Brauer. The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations. *European Journal of Social Psychology*, 32(6):853–867, 2002.

[6] Elizabeth Chell. Critical Incident Technique. *Essential Guide to Qualitative Methods in Organizational Research*, pages 45–60, 2004. https://doi.org/10.4135/9781446280119.n5.

[7] John M. Darley and Bibb Latané. Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4, Pt.1):377–383, 1968.

[8] Dominic DiFranzo, Samuel Hardman Taylor, Franccesca Kazerooni, Olivia D. Wherry, and Natalya N. Bazarova. *Upstanding by Design: Bystander Intervention in Cyberbullying*, page 1–12. Association for Computing Machinery, New York, NY, USA, 2018.

[9] Emile Durkheim. *The division of labor in society.* New York: Free Press, 1947.

[10] Michael Hechter and Karl-Dieter Opp. *Social Norms*. Russell Sage Foundation, 2005.

[11] Val Hooper and Tarika Kalidas. Acceptable and Unacceptable Behaviour on Social Networking Sites: A Study of the Behavioural Norms of Youth on Facebook. *The Electronic Journal Information Systems Evaluation*, 15(3):259–268, 2012. https://issuu.com/academic-conferences.org/docs/ejise-volume15-issue3-article830.

[12] Robert M Krauss, Jonathan L Freedman, and Morris Whitcup. Field and laboratory studies of littering. *Journal of Experimental Social Psychology*, 14(1):109–122, 1978. https://doi.org/10.1016/0022-1031(78)90064-1.

[13] Klaus Krippendorff. The Cybernetics of Design and the Design of Cybernetics. *Design Research Foundations*, pages 119–136, 2019.

[14] Bibb Latané. *The unresponsive bystander: Why doesn't he help?* Appleton-Century Crofts, 1970.

[15] Caitlin McLaughlin and Jessica Vitak. Norm evolution and violation on facebook. *New Media & Society*, 14(2):299–315, 2012. https://doi.org/10.1177/1461444811412712.

[16] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2009. http://www.sup.org/books/title/?id=8862.

[17] L. Peter, L. Diane, A. Marc, F. David, F. Betty, and J. E. McGrath. Bystander effect in a demand-without-threat situation. *Journal of Personality and Social Psychology*, 24(2):166–171, 1972. https://doi.org/10.1037/h0033380.

[18] Yasmeen Rashidi, Apu Kapadia, Christena Nippert-Eng, and Norman Makoto Su. "It's easier than causing confrontation": Sanctioning Strategies to Maintain Social Norms and Privacy on Social Media. *Proc. ACM Hum.-Comput. Interact*, 4(CSCW1, 23):1–25, 2020. https://doi.org/10.1145/3392827.

[19] William Ross Ashby. *An Introduction to Cybernetics*. London: Chapman & Hall, 1956.

[20] Reihan Salam. The Facebook Commandments, 09 2007.

[21] M. Sherif. *The psychology of social norms*. New York: Harper, 1936.

[22] Statista. Number of global social network users 2017-2025, Jan 2021. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/.

[23] Karen Verswijvel, Wannes Heirman, Kris Hardies, and Michel Walrave. Adolescents' reasons to unfriend on facebook. *Cyberpsychology, Behavior, and Social Networking*, 21(10):603–610, 2018. https://doi.org/10.1089/cyber.2018.0243.

[24] Norbert Wiener. *Cybernetics: Or Control and Communication in the Animal and the Machine*. Cambridge, Massachusetts: MIT Press, 1948.

[25] Wikipedia contributors. Bystander effect, Juli 2021. https://en.wikipedia.org/wiki/Bystander_effect.

[26] Wikipedia contributors". Cybernetics, 08 2021.

[27] Leping You and Yu-Hao Lee. The bystander effect in cyberbullying on social network sites: Anonymity, group size, and intervention intentions. *Telematics and Informatics*, 45:101284, 2019. https://doi.org/10.1016/j.tele.2019.101284.

[28] Alyson L. Young and Anabel Quan-Haase. Information revelation and internet privacy concerns on social network sites: A case study of facebook. In *Proceedings of the Fourth International Conference on Communities and Technologies*, C&T '09, page 265–274, New York, NY, USA, 2009. Association for Computing Machinery. https://doi.org/10.1145/1556460.1556499.

Appendix A

General interview structure

Starting from the next page, we have included the general structure that we have used for the conducted in-depth interviews. The interviews have been held in Dutch and therefore this interview structure was originally in Dutch as well. However, in order to improve the flow of this research paper we translated it entirely to English. The structure also contains certain comments shown in italics, in order to explain how and why certain questions have been asked in the way they have been asked. Please note that this structure was only used as an outline and main thread of the interview. Often, multiple follow-up questions were asked depending on the answers participants gave. This made the proceedings of each interview unique.

Main general structure / guideline of the conducted interviews:

1. How old are you?

2. How would you describe your gender?

3. Approximately how often per week do you use social media?

4. What kind of social media platforms do you then primarily use?

Now I will give an explanation about the goal of the research and ask questions about how one reacts to other people and online friends violating the privacy of others by sharing content without their consent.

5. How do you feel about such behavior? (Is this behavior acceptable/unacceptable?)

<u>Possibility 1 – Ask these questions if the following holds:</u>
We first ask whether the respondent is able to recall one or more moments that happened in the past where online 'friends' violated the privacy of others by sharing content or information without having their permission. In order to clarify and concretize the situation, we could give an example like an embarrassing picture being shared via Twitter. If the answer is yes and the respondent is able to name a situation, we can ask follow up questions in order to find out how and why he or she reacted at the time. If the respondent chose to not react, it is interesting to find out why not and whether the bystander effect could have played a role. These questions are formulated here:

6. Do you recall one or more moments that happened in the past where online 'friends' (people you speak or spoke to via social media) violated the privacy of others by sharing content or information about others without having the permission to do so? For example, consider sharing an embarrassing photo, like a nude photo, of someone on Twitter.

    6.1. If so, do you remember how you reacted at the time?

    6.2. Why did you respond in that way?
    (Even if you decided not to respond or intervene, why not? *Bystander Effect*?)

    6.3. If the responded decided not to respond or intervene:
    What could have prompted you to intervene in this situation anyway?

    6.4. How long did you know the person who shared this content?

    6.5. In retrospect, do you think you responded well? How do you think you should have responded and do you feel this would have been better?

Possibility 2 – Ask these questions if possibility 1 did not hold:

In the case that the respondent could not recall one or more moments that happened in the past where online 'friends' violated the privacy of others by sharing content or information without having their permission, we try to describe different situations and ask the respondent what they think they would do in these situations. By concretizing the situations as much as possible we try to make sure that respondents will describe what they really would do in a particular situation instead of explaining what they would want to do. Also, we try to minimize the effect of a privacy paradox. We also randomize the order in which the three different situations will be presented to the respondent, such that given answers to the first situations do not influence the answers given in the later situations in the final results.

It is important to make clear to the respondent that the described situations occur in a 'public space' on social media. This is because that is where the bystander effect might occur.

Situation outline A)

Please mention the first name of a very good acquaintance. Someone you regularly see in real life and also speak to via social media. Now, imagine that [name] shares a very embarrassing photo of someone, like a nude photo or a photo of someone being very drunk, via social media such as Twitter or in a WhatsApp group. Everyone on Twitter or in this WhatsApp group is able to see this photo and the person in the photo knows nothing about this. In result, his or her privacy is being violated by this.

Situation outline B)

Please mention the first name of someone you interact with very occasionally, e.g. a few times a year. Someone you added or started following years ago on social media like Twitter or Facebook. Now, imagine that [name] shares a very embarrassing photo of someone, like a nude photo or a photo of someone being very drunk, via social media such as Twitter or in a WhatsApp group. Everyone on Twitter or in this WhatsApp group is able to see this photo and the person in the photo knows nothing about this. In result, his or her privacy is being violated by this.

Situation outline C)

Please mention the first name of someone you added or started following on social media like Twitter, but haven't heard from in years. Now, imagine that [name] shares a very embarrassing photo of someone, like a nude photo or a photo of someone being very drunk, via social media such as Twitter or in a WhatsApp group. Everyone on Twitter or in this WhatsApp group is able to see this photo and the person in the photo knows nothing about this. In result, his or her privacy is being violated by this.

7. What would you do in this situation? Do you think you would react?

    7.1. If so, how would you react?

        7.1.1. Why would you respond in this way? What is your reasoning?

    7.2. If not, why would you not intervene in this situation? What is the reason for you? (*Maybe because nobody replied? -> Bystander effect?)*

        7.2.1. What could prompt you to intervene in this situation anyway? How would the situation need to change?

8. To what extent do you think the way you would respond depends on how good your relationship is with the person engaging in the behavior of violating another's privacy, and on how often you speak to someone?

9. Do you think that the fact that everyone can basically "read along" with what happens on social media has an effect on how you would handle this situation?

Appendix B

Compact overview results

Privacy and Good manners on the social web - Results				
			Indiciated by # participants	**Total number of participants**
Opinion about privacy breaching behaviour:	"Unacceptable"		14	14
	"It depends"		14	14
Reaction Types:				
	Privately:			
		Sending a direct message	12	15
		Sending a message via WhatsApp	3	15
		Calling	2	15
	Reluctant		13	15
	Publicly		4	15
	Blocking or removing		5	15
	Discuss with friends		4	15
	Other		2	15
Relationship with privacy violator:				
	Direct friends:			
		React in private	12	15
		React reluctant	3	15
		Intervention is easier	12	15
		Intervention is more complicated	3	15
	Acquaintances:			
		React in private	11	15
		Intervention is easier	12	15
		Intervention is more complicated	1	15
		Reporting them is less likely	2	15
	Strangers:			
		React reluctant	11	15
		React publicly	4	15
Severity of privacy violation:				
	Severity influences intervention		11	15
	Severity does not influence my intervention		1	15
Reporting mechanism:				
	Making use of the feature		13	15
	Never made use of the feature		2	15
Bystander effect:				
	No influence:		12	15
		Influence is limited	8	15
	Influence		3	15

Table B.1: *Compact overview of statistics of the results*