Thèse
de doctorat
de l'UTT

Malte PASKUDA

The Influence of Anonymity on Participation in Online Communities

Spécialité :
Ingénierie Sociotechnique des Connaissances, des Réseaux et du Développement Durable

2016TROY0033 **Année 2016**

utt
université de technologie
Troyes

THESE

pour l'obtention du grade de

DOCTEUR de l'UNIVERSITE DE TECHNOLOGIE DE TROYES

Spécialité : INGENIERIE SOCIOTECHNIQUE DES CONNAISSANCES, DES RESEAUX ET DU DEVELOPPEMENT DURABLE

présentée et soutenue par

Malte PASKUDA

le 24 octobre 2016

The Influence of Anonymity on Participation in Online Communities

JURY

M. M. BAKER	DIRECTEUR DE RECHERCHE CNRS	Président (Rapporteur)
Mme N. GAUDUCHEAU	MAITRE DE CONFERENCES	Examinateur
Mme M. LEWKOWICZ	PROFESSEUR DES UNIVERSITES	Directeur de thèse
M. M. PRILLA	PROFESSOR	Rapporteur
M. M. ROHDE	DOKTOR	Examinateur

Acknowledgements

Myriam Lewkowicz

Michael Baker

Michael Prilla

Nadia Gauducheau

Markus Rohde

Michel Marcoccia

Valentin Berthou

Matthieu Tixier

Hassan Atifi

Ines Di Loreto

Karine Lan

Lorraine Tosi

Aurlien Bruel

Khuloud Abou Amsha

Josslyn Beltran Madrigal

Les membres de lquipe Tech-CICO

et les membres du projet TOPIC

trouvent ici mes remerciements les plus sincres.

Abstract

This work presents my PhD thesis about the influence of anonymity on participation in online environments. The starting point of this research was the observation of the design process of an online platform for informal caregivers. I realized that there is no knowledge about the practical effects that an anonymous identity system would have. This thesis contains the subsequent literature review, which has been synthesized into a model that shows which participation factors might be influenced by anonymity. Three studies on existing online environments have been conducted: One on Youtube, where there was a change in the comment system forbidding anonymous comments; one on Quora, where users can choose to answer questions anonymously; and one on Hacker News, where users choose how many identity factors they want to present and which name they use. The results of these studies are that, contrary to what the literature would suggest, 1) anonymity did not result in impolite and uncivil discussions, and 2) other factors than anonymity have a stronger influence on participation, which means that 3) anonymity can make the effect of social signals visible, e.g. text properties like length influencing social appreciation. Additionally, it appears that participation is linked to profile completeness, and that an established web presence elsewhere limits participation. The implications of these results are a confirmation of the Social Identity Model of Deindividuation Effects, according to which anonymity can have positive effects on group identity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Problem description

Researchers in Human Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW) have been investigating how people interact with each other using technologies for many years. Sometimes this research reaches into areas where other disciplines, but also people in general, have already made their own opinion. One such area is anonymity.

Anonymity is often believed to be a negative characteristic of online environments. The anonymity that a user can achieve online is believed to be responsible for the negativity of online discussion, going so far as it being considered as a "danger for society". Political will to ban anonymity has been strong: Politicians tried to enforce registration schemes for internet usage, including measures on the form of each and every comment that could be found online[1]. Historic psychological theories like the Deindividuation theory focus on

[1] See https://www.taz.de/!5274217/. The politician Fischer is arguing for a complete ban of being anonymous online, as according to him, the quality of discussions suffers from being anonymous. He said that anonymity determines whether a user thinks he is responsible or not for his statements:

> Anders sah das der damalige Vorsitzende der Enquete-Kommission "Internet und digitale Gesellschaft" Axel Fischer (CDU), der sich fr ein "Vermummungsverbot im Internet" aussprach. Fischer argumentierte, dass unter der Mglichkeit sich pseudonymisiert im Netz zu uern "die Qualitt von Diskussionen in Foren und Blogs" leide. Die Anonymitt verleite Nutzer zu uerungen, die sie hinterher bereuen knnten. Er halte es fr bedenklich, dass sich Nutzer durch ein selbst gewhltes Pseudonym vermeintlich jeglicher Verantwortung fr uerungen entzogen.

how anonymity in groups leads to mob behaviour, that is on how a group of, otherwise normal, people can transform into a raging, obscene and violent mess (see section 2.2.1 on page 12).

However, the question of how to treat anonymity is not one that can be easily answered by looking at prior research. This is what we observed in our research, which aims at designing a new platform for social support among old informal caregivers (TOPIC). Despite all the negative hypotheses about general anonymity online, there are clear advantages in allowing an anonymous identity model for a social support platform: A user would be expected to be more willing to actually use the platform to share the more intimate aspects of being a caregiver. For example, users could more openly talk about the degrading and humiliating aspects of having to manage / linked to the changes in behavior of a person, suffering for example from Alzheimer's Disease (Salem et al., 1997). Also, in a more general context, is the option of being anonymous online not an important tool to construct its own online identity (Nagel and Frith, 2015)?

Anonymity would therefore be linked to effective use of the platform. This is an important issue since the willingness to use the platform equates to a raise in participation. Raising participation is important for most platforms, but especially for platforms whose aim is to provide social support among peers: If the designed platform is not used, the support cannot take place. In the case of TOPIC, the goal of the platform was to help informal caregivers by providing a place for them to exchange social support, in all three dimensions: Informational, emotional and tangible. It was thus crucial for the platform to succeed in motivating the caregivers, who often do not have much time and resources for additional activity.

What are the critical factors that influence participation, experience sharing and the general user satisfaction for an online collaborative systems? This is difficult to know beforehand. For a platform like TOPIC, it is crucial that people participate: The more they participate, the higher the chance that they get support themselves. Experience sharing is absolutely important, as it is a key factor in giving and getting social support (Salem et al., 1997).

Therefore, the research question is: what is the influence of anonymity on user participation in general, and on experience sharing in particular? More specifically, what would be the best option when designing a tool for a community like the one under concern in the TOPIC project? Should the user be

able to choose to stay anonymous, choose his online user name, or be forced to use his real name? It is also a very generalizable practical question: When designing an IT-system, should the user be able to stay anonymous, choose his online user name, or be forced to use his real name?

Based on these questions, the research work that I am presenting here investigates how anonymity influences participation. What is meant exactly by anonymity and participation?

Anonymity in this context means writing online under a name that is not linked to one's own civil identity, including having no name at all. It can thus also mean writing under a pseudonym. This is possible when subscribing to the assumption that identity is dividable, that there are different aspects in each identity that one can choose to present or not. This issue of identity would therefore influence one's behaviour.

Participation refers to writing posts on a platform. The online communities that have been examined in this research are text-based, and the direct way of participating is writing a new entry or comment. This also holds for the envisioned care support platform. There are peripheral means of participation which will be mentioned where applicable, but they were not the focus in my research.

There is prior research in this area that are of interest here. Participation has been investigated under the perspective of finding universal and specific factors that favour or inhibit participation in online communities. The focus has, therefore, been more on motivation. Anonymity has been examined in several fields like Social Psychology, and its effect has been studied both online and offline. There are also several important theories that try to explain its effect.

But despite this existing large theoretical background, there are only few studies that make the link between anonymity and participation and look at the real effects of what arose in the online wild. Moreover, those studies contradict each other, resulting in this current situation: HCI cannot inform designers of new online systems about the consequences of allowing anonymity or not.

This work tries to improve this current situation. The research scheme I followed is that, first, I looked at the existing literature to understand the current knowledge of my field. Second, I synthesized this knowledge into a model that allowed me to formulate hypotheses about the possible effect of anonymity online.

Third, we made studies on real online environments to test our hypotheses and generate empirical knowledge about how anonymity influences participation.

The first result is that anonymity does not automatically lead to impolite and uncivil discussions.

However, existing prior research has shown that anonymity inherently has a negative influence. Examining this hypothesis in the case of TOPIC, a second result is: The effect of anonymity can be overridden by other factors, for example by community culture and user interface. For instance, we observed this phenomenon on Youtube with the effect of the comment system switch, where the non-anonymous comments did not conform to expectations (see section 4.1 on page 45).

In reverse, this means that anonymity can uncover effects that would, otherwise, be masked. The third result is that we observed that social signals can override the effect of text properties and how anonymity can prevent this. I.e, in our study, social appreciation is directly correlated with comment length. And this applies only to anonymous comments.

Two additional results are that participation is linked to profile completeness, and last, that an established web presence that is elsewhere limits participation on other platforms.

This dissertation presents my work in the next chapters as follows:

1. Chapter 2 is the state of the art. It presents the literature that is relevant for our the research question. It starts by presenting the theories and perspectives of anonymity across several research fields. Then, it shows the existing studies and results on anonymity. Finally it presents the literature on participation.

2. Chapter 3 explains, in more detail than this introduction, the scientific approach of this work. It explains how the model was devised and which factors it contains, detailing, for each factor, the reason why it was included.

3. Chapter 4 gives a detailed presentation of the studies we made using the model that we developed in this research. We looked at the impact of anonymity on participation on three existing online platforms - Youtube, Quora and Hacker News - with each study further developing the results of the prior one.

4. Finally, chapter 5 concludes this thesis by summarizing its results. It also gives an outlook on future work, that could and should be done based on this work. Indeed, we believe that there are several ways to further test the developed hypotheses and to design fitting systems.

Chapter 2

Literature Review

2.1 Introduction

The literature review done for this thesis looked at anonymity across several disciplines. Just as Human-Computer-Interaction (HCI) is an interdisciplinary field, the question of the influence of anonymity is one that cannot be associated to only one discipline. I looked mainly at work published in the HCI and CSCW domains and at those attributed to Social Psychology, which already spans a broad area. But additional influence comes from other fields, for example the perspective of political and journalistic science, which looked at the question under the aspect of the public discourse or just the practical implication of a comment section.

Interested in the relevance rather than the discipline of a publication, I searched for publications that describe anonymity, pseudonymity - in some cases as an aspect of identity as a whole - and participation.

This section will lay out the most prominent positions, theories and results found in this studied literature.

2.2 Anonymity and Identity

Since anonymity and identity are subjects that have been largely studied, I will start by describing the central theories that are used to address this topic. The next section (see section 2.2.2 on page 16) will show interesting studies that

were conducted and their results.

2.2.1 Theories and Perspectives

There are two central theories usually used to explain the effect of anonymity on identity, and, thus, its effect on behavior.

The first one is classical: **Deindividuation theory** (Postmes et al., 2002), which goes back to theories about the functioning of groups from the 19th century (Le Bon, 1896; Reicher et al., 1995). Its main idea is that the individual norms of one person get lost when that person is in a large-enough group. Through anonymity and loss of personal responsibility, the single person in the group reverts to primitive and hedonistic behaviour, resulting in typical mob behaviour.

The second theory is the **social identity model of deindividuation effect** (SIDE) (Reicher et al., 1995). SIDE can also be seen as a deindividuation theory, but it explains the mechanisms and outcomes of anonymity in groups differently (Cress, 2005). According to SIDE, members of a group do not only lose their social norms, but they adapt to the norms of the group. Those norms can be in conflict with societal norms, but it is not a reversal to a primitive normless state. Also, anonymity works differently in this model: it minimizes the differences between the other individuals, allowing a higher identification to the group. Anonymity strengthening group identification can then equally result in behaviour varying from societal norms, but it also favors the group's norms, which does not necessarily have to be negative.

We have to be aware here that those theories differ in their underlying assumption of what humans are. The reversal theory of Le Bon assumes that humans are primitive beings, and that their primitive behavior is overridden by societal norms. Rousseau would disagree, and maybe he could follow the explanation of SIDE theory. SIDE theory also allows to defend the existing position of egalitarian groups as the anarchist hacker movement (May, 1992): anonymity is a good thing for groups and discussions, as the removal of status symbols removes differences, allowing a stronger identification to the group and a greater focus on what is said, therefore changing how the discourse process works.

While SIDE and deindividuation theory are the most prominent ones, there are several other theories and perspectives on anonymity and the web.

Assuming that online exchanges are influenced by anonymity and identity, we assume that the online space is not a space that is only dedicated to sharing and storing knowledge. To the contrary, it implies a vision of the web as a social space which is heavily influenced by emotion and social signals. These are transported even if text remains the main communication medium (Derks et al., 2008).

The role and impact of anonymity can also be discussed as an ethical question (Bodle, 2013), instead of looking at it under the aspects of its functional impact. Bodle argues that:

> ... anonymity in networked digital communications is indispensable as an enabler of other inalienable rights including informational privacy and freedom of expression. (ibid., p. 22)

Bodle describes a conflict between the positive impact that anonymity can have, with regards to freedom of expression and privacy functions. His perception is that of an industry moving to persistent online identity. Whether one follows the argumentation or not, in any case the paper contains an interesting list of current systems taking measures against anonymous accounts and access, like Facebook and Google+.

The opposing perspective is that anonymity achieves unaccountability, resulting in a lawless network and thus a lawless society (Davenport, 2002). This effect was also described as incivility and as a breakdown of the public sphere (Santana, 2012). It is my impression that research in HCI and CSCW is influenced by that negative perspective; We can indeed note that the discussion about technical identity management systems includes the idea that full anonymity is harmful, and that instead *obligation management* is needed (Borcea-Pfitzmann et al., 2006).

Both perspectives regard anonymity as an effect on users, without looking at the anonymity of our communication infrastructure. Edman and Yener (2009) have listed several ways in which anonymous communication systems are currently constructed, and how they are attacked. An anonymous communication system, according to their definition, is not one that simply avoids showing the user's name on a webpage, it is one that does not allow observers of the network traffic (like the NSA) to find out who transmitted or received which information. As more users realize the non-anonymity of all unprotected Internet usage, ex-

isting attitudes of users could change: Where users felt anonymous before, they could now be aware of the agency observers. This gets more severe with the emergence of more sophisticated attacks for determining the existing relationship in superficially anonymous looking networks (Shmatikov and M.-H. Wang, 2006).

But even without looking at the network, it was hypothesized that different kinds of anonymity are used in different kinds of studies, and thus of online systems (McLeod, 1997). As described there, the gap between the technical and the social aspects of anonymity were defined in the HCI-literature only in the early 90s. Technical anonymity is defined as the technical means to reach anonymity, like suppressing the user name (at that time, the discussion did not include the fact that the network was not safe), whereas the social side is whether the user feels anonymous. From my perspective, it clearly appears that user behavior is mainly controlled by the social aspect, which is why I will focus on that.

This relevant distinction was described by Valacich et al. (1992). The authors start by distinguishing different types of anonymity, specifically giving a definition for context anonymity (*the extent to which group members can identify the source of a particular contribution by recognizing the author through an identifier embedded in the contribution*) and process anonymity (*the extent to which group members can determine who is participating by directly observing who is making a contribution*)

(McLeod, 1997) introduces upon that distinction a model for the different aspects of anonymity, combining Mechanism, Attribution Direction, and Domain. Attribution Direction has not been discussed here already. It is the simple idea that there is a difference between oneself being anonymous, and the others being anonymous for oneself (see fig. 2.1 on the next page). In his model, attribution direction is at the left side of the cube. At the front, there is the mechanism, and social at the bottom:

> Source dissociation refers to a feeling that others cannot identify one as the source of specific messages, while identitylessness is associated with feeling that others dont know that one is a participant or what ones role in a session might be. (ibid., p. 224)

Then, there are the technical factors above, with content and process anonymity,

as described in the paragraph above. At the top of the cube model is the domain:

> The second dimension is concerned with the domain of anonymity and can be divided into message source and participant presence. The message source category refers to the ability to attribute specific messages to a specific source. ... The participant presence category refers to knowledge about the presence of other group members. This includes knowing the number of group members, knowing whether a particular individual is a member, knowing general characteristics of the other group members, and knowing how much any one group member participates. (McLeod, 1997, p. 225)

Dividing anonymity many different parts as McLeod does, can serve as a means to distinguish between different types of anonymity. This can be useful in different use cases.

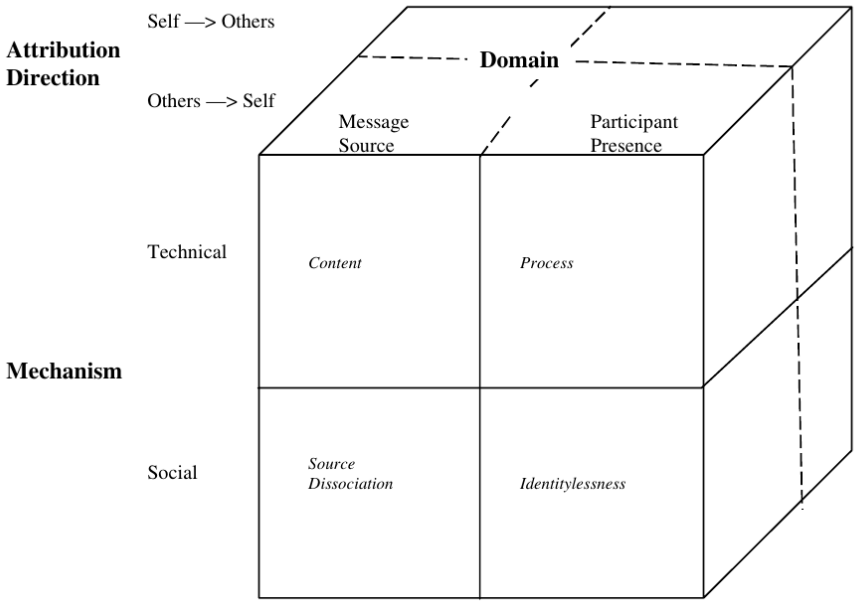

Figure 2.1: Proposed three-dimensional anonymity model by McLeod, taken from McLeod (1997).

The effect that people behave differently online was also called the *online disinhibition effect* (Suler, 2004). Apart from the focus on anonymity, as SIDE

and Deindividuation theory provide, there exist several alternative explanation models in which anonymity is just one factor. One is that online communication allows the participants to present a different aspect of their identity. This does not have to be the "true self"; the identity the user strives to achieve (which does not have to be positive, i.e. showing repressed anger). Suler argues that six factors of online communication can cause a shift:

> The disinhibition effect can then be understood as the person shifting, while online, to an intrapsychic constellation that may be, in varying degrees, dissociated from the in-person constellation, with inhibiting guilt, anxiety, and related affects as features of the in-person self but not as part of that online self. (Suler, 2004, p. 325)

All those theories could depend on the actual user motivation to be anonymous. User interviews showed that there are many different motivations, and also different ways in which users try to achieve their goal (Kang, Brown, et al., 2013).

In mentioning user interviews at this point, my objective in the argumentation is to move away from general perspectives and theories, which are not based on any empirical data, to studies and experiments. Those will be described in section 2.2.2.

2.2.2 Studies on Anonymity

Regardless of which theory and perspective one subscribes to, there are several studies that have looked at the actual effect of anonymity, that one can try to use to argue for or against specific theories. A lot of these studies are laboratory experiments that may or may not be relevant to real online environments. But some studies also looked at real existing online environments.

A study that appears central for us is the one conducted by Kilner and Hoadley (2005). They were able to observe the stepwise move of an existing forum for US-soldiers, from a user account model allowing anonymous users, to another one using their civil identity, or in that case, their military identity. They were able to measure the impact of that change by looking at participation, like numbers of thread views and the number of comments, and at the quality of the comments, via a custom coding scheme. The authors saw a strong decrease

of negative comments and an increase in participation as the experiment reached the level of pseudonyms.

A comparable case study was done when the tech-news site TechCrunch changed from the comment system Disqus to Facebook comments, moving from anonymous comments to forbidding them (Omernick and Sood, 2013). Omernick and Sood had a dataset of around one year each, from before and after the switch. They analysed them *through measures of reading level, relevance to the target article, negativity and presence of swear words and anger words.* That was combined with an analysis of participation measurements. Their results saw a better quality in comments (less swear words, higher relevance) when commenters revealed more of their identity, and they additionally saw a gap between how much anonymous and pseudonymous Disqus comments were liked. Their look at how the amount of participation was affected was inconclusive, showing decreases in some and increases in other areas.

A more recent case study is the analysis of comments seeking for social support on reddit (Andalibi et al., 2016). Andalibi et al. looked at comments in subreddits providing support for victims of sexual abuse. In those subreddits, people can use whatever name they want, and it is common to use a pseudonym. But these pseudonyms are not necessarily anonymous, e.g. if they wrote comments before that allow others to deduce their real civil identity. Andalibi et al. thus looked at the use of throwaway accounts, accounts that are generated for the single purpose of writing one specific comment anonymously. These accounts often contain throwaway in their username, or they write about being a throwaway. It was observed that these accounts were often used to ask for social support. Their usage also coincided with significant linguistic differences, something that is noteworthy as it could serve as a means to detect the perceived degree of anonymity. The differences mean that one can take posts on reddit and use the classifier that is generated, based on word choices alone, to see whether the user thinks he uses a throwaway account. By extension, this could mean that though looking at those differences also allows to see whether the user thinks he is anonymous.

These studies stand at the front of this section, but they are not the most common. Historically, a wide range of studies tried to detect the effect of anonymity by conducting lab experiments.

One recent and often cited publication of one such experiment is Lapidot-

Lefler and Barak (2012). Lapidot-Lefler and Barak confronted 142 participants with a dilemma they should discuss and solve in an online chat system. They were either anonymous or not, invisible or not and had eye contact or not. The goal was to see which variable had the biggest effect on social disinhibition, i.e. flaming. Their conclusion was that anonymity and invisibility - the factors such behaviour is usually attributed to - did not result in more flaming. It was the lack of eye contact which had the biggest effect. Their study raises the important question of when exactly a user in an online environment feels anonymous.

Not being visible to other participants seems to be a big aspect. In an experiment where computer mediated communication was already seen to provide higher levels of self-disclosure, the moment a video image of the participant was shown, the amount of self-disclosure lowered significantly. A subsequent experiment concluded that the effect stems from the changed value in self-awareness (Joinson, 2001). But it was seen that, in practice, users who were not writing under their real name (being more anonymous) shared less information, while those showing a photo of themselves provided more (Hollenbaugh and Everett, 2013) - a hint that the latter study mixed up cause and effect. Still, common result is that anonymity can increase self disclosure (Hollenbaugh and Everett, 2013; Joinson, 2001; Ma et al., 2016).

The other side of this perspective is how identity is established in anonymous systems, which was observed on 4chan (Bernstein et al., 2011). Bernstein et al. saw that while 90% of posts on 4chan were made fully anonymous, those anonymous posts sometimes used alternative identity signals like timestamps (pictures of handwritten notes with the current date), the use of slang and the crafting of special symbols like a triforce.

The effect of anonymous communication via the Internet on expression of the actual vs the true self (Bargh et al., 2002) was also tested in a laboratory experiment. Those concepts of self need to be explained first. They are based on the idea that every person has several versions of his own self that are used in different contexts, going back to theories of Jung and Goffman in the 1950s (Goffman, 1949; Jung, 1953). The actual self is the version currently presented. The true self is the real inner self, but aspects of oneself that are not necessarily shown. It is clear that those are highly speculative and abstract concepts that may have no grounding in how people really behave. However,

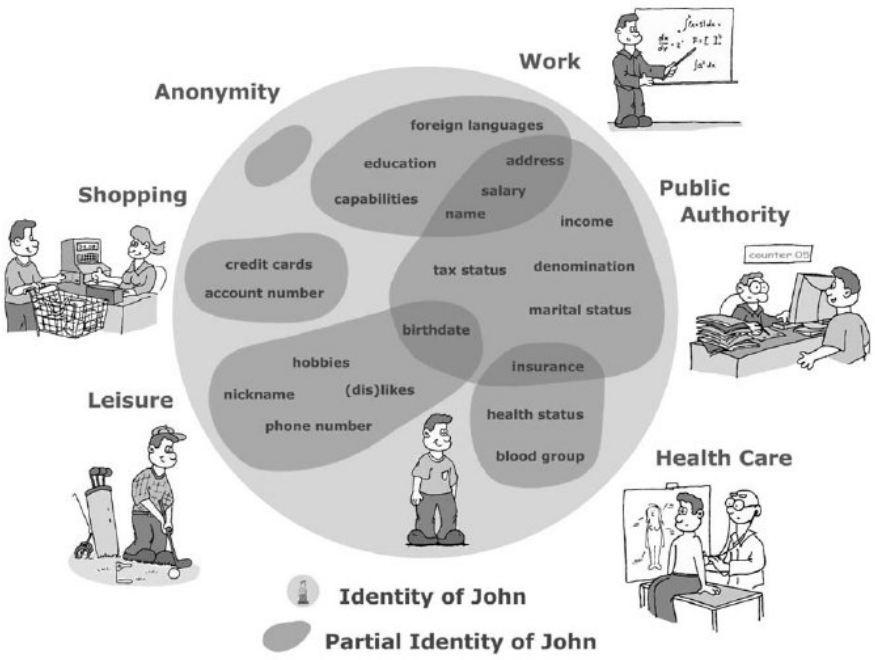

Figure 2.2: Depicting context-sensitive parts of an identity, taken from Borcea-Pfitzmann et al. (2006).

the idea that anonymity in general and anonymity on the Internet especially allows self-expression of facets of people's self that are not normally possible is a common idea. A variation of that idea is found in more technical-inclined modelizations of user identity, realizing that not all information about a user fits every context (see fig. 2.2 on page 19). Bargh et al. tried to test it in a laboratory via a Me/Not-me test, a test in which adjectives were asked to be assigned or refused, via two buttons, as part of the own self. The reaction time was measured. Those were then compared to the ones that the participant had declared before, as adjectives being part of his existing abilities and characteristics (actual self) and those he would like to possess (true self). The result was that reaction times for concepts of the actual self were faster.

Studies like the ones presented above can be easily criticized. What is measured there relies on interpretation of reaction times, and the concepts used - the true and actual self - could also be taken from a religious text. From our point of view, these kind of results are much less interesting than results of real interaction in a real environment.

Still, tests like this can uncover interesting aspects, like in the third experiment by Bargh et al.: After having self-described how they see each other, participants also described what they would like to see in a friend and in a romantic partner. They then talked with each other, either in a room or via a chat room. The result was that the people talking via the chat room liked each other a lot more. The partners of the Internet-based communication were more able to convince the others that they had the qualities of their true self. It is one of the clear example showing how anonymity or the omission of social signals can improve the judgement of another person.

To think about this result as the presentation of a true self is, of course, not imperative. Alternative explanations are easy to find, as the above presented explanation of Suler for the online disinhibition effect.

Laboratory experiments where also used to investigate the impact on source credibility. When presented with an argument, Rains (2007) tested whether the argument from an anonymous user is more or less convincing. In an experimental task, the participants were asked to make a decision regarding a dishonest team leader. The confederate presenting a solution was, in some cases, anonymous. When he was anonymous, participants thought he was less trustworthy, less persuasive and was believed to have less goodwill towards the group.

At this point, one could have the impression that anonymity is regarded as a negative factor for online systems. But there is no consensus about that. Several studies observed the use of anonymous systems and found them helpful and positive. In Nelimarkka et al. (2014), Nelimarkka et al. looked at the usage of an anonymous backchannel for teenage pupils during class. They were careful to give this precision, the fact that in small school classes full anonymity is unlikely to exist, but their system provided no identity cues whatsoever. Their system providing the backchannel was a chat system, allowing to ask questions and to comment on them. Given the impression of anonymity painted so far, we could expect that the system would have been used for insults and interpersonal attacks. But that was not the case; the pupils evaluated the use of the system as respectful, anonymity as agreeable, and the system overall as helpful.

There are also more conflicted results. It was observed that electronic groups reach better solutions in social dilemmas, but are subsequently less able to implement those solutions (Rocco and Warglien, 1996). The members of the electronic groups were not anonymous, but Rocco and Warglien theorized that the breakdown in cooperation was caused by the inability of the participants to lead an ordered conversation via the email system that was being used, inhibiting the emergence of a group identity. This fits nicely to the result of a study done with wikipedia authors, that showed that anonymity slightly prevents group-conformity (Tsikerdekis, 2013).

Researchers also realized that different online tools provide different degrees of anonymity/privacy, and that those tools are not equally suited for different tasks. These tools provide, for example, different degrees of social presence, which can hinder their usage (Tu, 2002). Users were willing to forgo privacy if they gained better social presence.

The fact that usability trumps privacy concerns was also stated as an explanation of why more aggressive chat applications and social networks, like Whatsapp and Facebook, won against their more careful competitors. When asked about the the violations of privacy as risky, users have the tendency to depict a more complicated situation, but to not object to them when they occur (Phelan et al., 2016). This observed behaviour is a challenge for designers of systems respecting privacy.

An interesting model of combined anonymous and non-anonymous communication, that could possibly combine the positive aspects of both, was described

by Birnholtz et al. In Facebook confession boards, users send anonymous posts to moderators that post them eventually into groups. The user of these groups can then discuss under these posts using their civil identity. In their case study, Birnholtz et al. saw very little negativity in the responses and very open questions (Birnholtz et al., 2015), which indicates that combined models can work well.

A block of research literature is dedicated to examining users who do not participate directly in online communities, depending on the ideology called lurking or social-loafing. In parts of that research, anonymity is identified as a factor influencing participation (see section 2.3 on page 24). E.g. in Shiue et al. (2010), anonymity is seen as a positive factor:

> There is a general agreement that promoting freedom of expression and enabling the free flow of information are attributable to anonymity. The result indicated that anonymity will probably result in strong social ties.

Note though the qualifier *probably*. This stems from their research methodology, i.e. asking users of online communities via a survey how they feel about being impacted by several factors, which is probably not the most appropriate methodology to answer such hypotheses definitely.

A current phenomena of online anonymity is the raise of anonymous communication apps and the research focusing on them. Anonymous communication apps hereby means software targeting smartphones, that allows groups of people to communicate without having identifiers or fixed topics. Examples are Whisper (see fig. 2.3 on the next page), YikYak, and Secret.

A difference to existing anonymous online boards like 4chan is the incorporation of location features, fostering interaction between people living closeby (G. Wang et al., 2014). Users profit from these apps because they think that anonymous communication allows more honesty, openness, and diversity of opinion, and their behaviour on those platforms is governed by behavior rules respecting privacy (Kang, Dabbish, et al., 2016). Fitting to common identity and SIDE theory, users described that they were feeling an attachment to the group, though how reliable this attachment is, was questioned by the authors. In accordance with prior research, users relied on alternative identity signals, like the language used and location, to make an opinion about the identity of other users. Ad-

Figure 2.3: The nearby tab in Whisper, as shown in Kang, Dabbish, et al. (2016).

ditionally, Kang, Dabbish, et al. described anonymity as a participation factor. This will be discussed again below as a small part of the following section 2.3 on the literature about participation.

2.3 Participation

As mentioned above, there is some literature dedicated to lurkers. Lurker is a denomination for people who are part of a community but do not participate, e.g. the silent registered reader of a bulletin board. The research of this behavior slides into the question of what motivates people to participate in an online community, which is normally defined as writing entries, asking questions and responding. But this can vary based on the type of the online community.

Sometimes, the literature uses the pejorative term "social loafing" instead of lurking. Shiue et al. (2010) for example uses that term, and writes:

> Previous research has established that knowledge sharing intentions are based on group cohesion. Several studies also suggested that social loafing behavior will seriously corrode group cohesion. Therefore, social loafing is a key obstacle to fostering online community development. (ibid., p. 768)

The research is thus often focused on activating lurkers, to foster participation. But the negative judgement of lurking is challenged by parts of the research. Nonneeke and Jenny Preece write:

> It is unfortunate that the term lurker, with all of its negative connotation, has gained acceptance. Fortunately, lurking can now be understood as the many activities related to membership in online groups. Rather than being free-riders, lurkers should be called participants (publicly silent though they may often be). (Nonneeke and Jenny Preece, 2000, p. 7)

There are several models on what motivates lurking, and motivation is in itself a research area. Shiue et al. researched a model in which lurking is influenced by social ties and perceived risks. These are themselves influenced respectively by offline activities and anonymity, and by media richness and knowledge quality (see fig. 2.4 on the next page). These influences incorporate the common idea

that motivation is influenced by the characteristics of the system and external social factors.

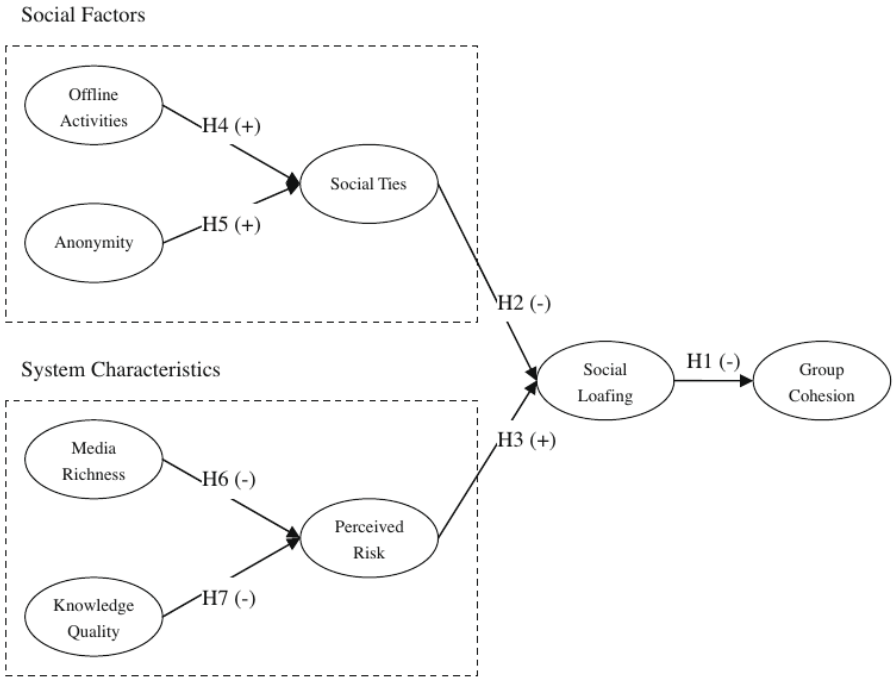

Figure 2.4: Research model of what motivates lurking, from Shiue et al. (2010).

How many users can be expected to stay passive varies on the topic and traffic level of a community (Nonneeke and Jenny Preece, 2000).

Whether people tend to not participate, either as part of their personality or whether that is defined by the characteristics of the system, is a discussion point in the literature. Since there are many users who are lurkers in some communities are active in others, it is likely that it is a mix, and that the tendency to not participate is a personal trait, but is also governed by the characteristics in questions (Muller, 2012).

This assumption is more likely to be correct given the prior research on participation factors, which is not specific to lurkers. It was seen in experiments that people confronted with moderated communities report a higher motivation to participate in these communities. The interactivity of messages (referencing prior messages) influenced that motivation as well when the response rate

was slow: if messages were more interactive, they reported higher motivation to participate (Wise et al., 2006). It was seen that transporting negative emotions boosted activity in the BBC forums (Chmiel et al., 2011). And specific rhetoric strategies used when making requests make it more likely to get responses (Burke, Joyce, et al., 2007). Knowledge of these factors can be used to try to design systems with the capability to persuade users to perform a specific activity (e.g. (Schneider et al., 2016), (Chang et al., 2016)).

Instead of focusing on specific system factors, the *common identity theory* and *common bond theory* as used by Ren, R. Kraut, and Kiesler give a powerful but easy to understand model on what drives user to participate in specific communities. Common identity theory covers the aspects the user likes about the group as a whole. Common bond refers to the users in the group with whom a user might have formed a connection. If users identify with the group as a whole, they will have less issues when individual members leave the group, than if they are bound to specific members of the group (Ren, R. Kraut, and Kiesler, 2007). Ren, R. Kraut, and Kiesler claim that one can design online groups in ways that strengthens identity or bond-based attachment to a group.

For this design Ren, R. Kraut, and Kiesler (ibid.) identified, via a literature review, three factors for each category that influences attachment to a group. They saw that social categorization, in-group interdependence and intergroup comparisons cause identity-based attachment, while social interaction, personal information and interpersonal similarity cause bond-based attachment (see fig. 2.5 on the following page).

Ren, R. Kraut, and Kiesler (ibid.) also state several design implications that, based on their theory, should influence identity or bond-based attachment. These cover the treatment of newcomers, the tolerance of off-topic discussions, the ideal community size, the role of core members, and the existence of subgroups.

The treatment of newcomers relates to a concept called eternal September, when an influx of new users endangers an existing community. While it was observed how communities were destroyed by that, it is now known that communities can survive a high influx of new users. Suggested factors are active and well-functioning moderators, an existing strong sense of community and tool-supported moderation (Kiene et al., 2016).

Another view on participation is to look at who already uses which online

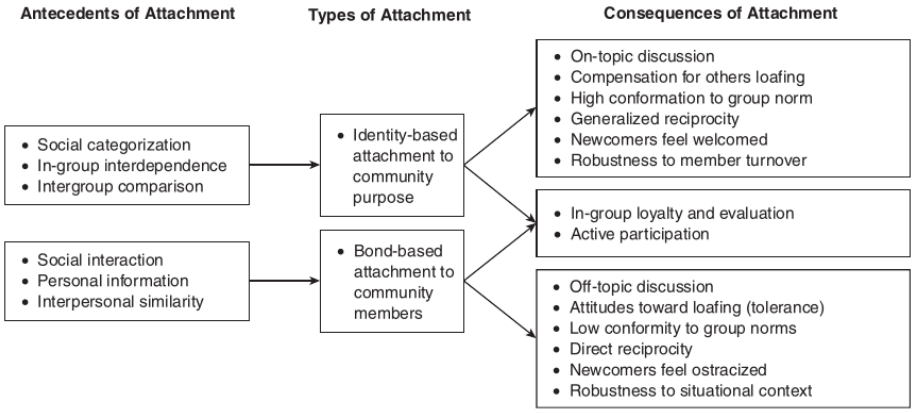

Figure 2.5: Factors influencing attachment separated into common identity and bond model and their consequences, taken from Ren, R. Kraut, and Kiesler (2007).

tools. For example, Harley and Fitzpatrick performed two case studies with the explicit goal of investigating the use of social networks by old people (Harley and Fitzpatrick, 2012), arguing that social networks can help by generating social connections and opportunities. Better tool-support is also a recent approach to help moderators identify good comments by using multiple scores and to present them visually (Park et al., 2016). This helped to manage participation and is also expected to raise it if those efforts result in higher quality discussions.

One perspective is to look at the technical design of the online platform of the community, in order to examine specific functions. That is done in the Community Activity Framework (Oostendorp and Varik, 2011) (see fig. 2.6 on the next page).

The frameworks contains a mix of functions that relate to factors already described here, like having photos in profiles, and other functions that were not discussed in other terms so far. Those are, for example, the use of graphical emotions in posts and having rules to guide the group discussions. Others are strictly functional rather than content based, like the availability of email notifications, post counts next to the post (that could also lead to social categorization and social interaction) or the availability of a related news section.

Sometimes, online participation is influenced by offline activities (Lpez and Farzan, 2015). Lpez and Farzan also noted that particular kinds of requests in the studied local forums are effective in order to generate online interaction,

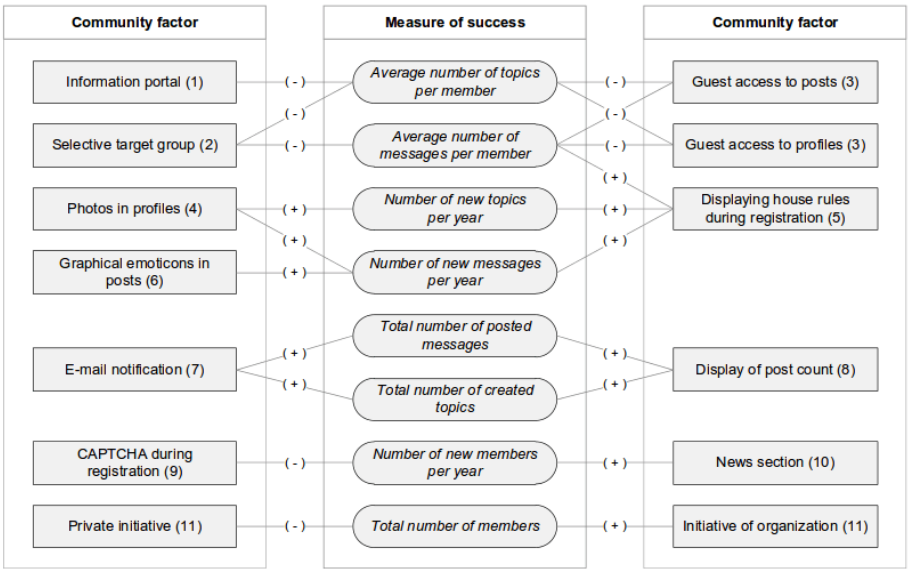

Figure 2.6: Factors described in the Community Activity Framework, taken from Oostendorp and Varik (2011)

but do not result in those users staying active on the platform. They state that, in this kind of local forums, online interaction (and thus participation) may be less important than in other online environments. This highlights the idea that the definition of successful participation can vary depending on the environment. A further spin on this perspective is the idea of raising social awareness about limiting participation, transforming participation into the right kind of participation, or managing non-participation (Ko et al., 2016).

Theories about motivation strongly influence the current research on participation done in HCI. In a study about the response to requests for help targeting friends on social networks ("Friendsourcing"), Zhu, Das, et al. looked at the effect of extrinsic motivation. In theory, extrinsic motivation stands in contrast to intrinsic motivation, the former being motivation that stems from additional factors (like money), and the latter the existing inner motivation to do something. Intrinsic and thus overall motivation can suffer when weaker extrinsic motivation is added. In this study, however, large monetary rewards raised participation and the author claimed that they also served as a scapegoat to preserve the image of having stable relationships when the requested help did not arrive (Zhu, Das, et al., 2016).

Motivation to participate and to react to changes in existing environments is the topic of a recent study focusing on the impact of policies on online communities by Centivany and Glushko. They reference political scientist Albert Hirschman and his exit, voice and loyalty framework. The framework contains, with exit and voice, two options that consumers of deteriorating services (and communities) have, and with loyalty, a strategy for services to react. Loyalty favours the voicing of critique over the exit reaction (Centivany and Glushko, 2016). The framework thus describes loyalty as a participation factor.

The focus of their study was a conflict on reddit over the unexplained firing of an organiser of a particular important subreddit, /r/IAmA, a place to ask willing participants (including prominent figures like US-President Barack Obama) all possible kinds of questions. The conflict transformed into a bigger conflict about the treatment of moderators, transparency and free speech, and uncovered a neglect of the technical base of the platform. During the conflict, many other subreddits temporarily shut down, forcing the then-CEO of the company behind reddit - already unpopular before for her perceived distance to the community and weakened by a controversy over a lawsuit with a prior employer over feminist positions - to step down. Centivany and Glushko (ibid.) contains a more detailed explanation, and Matias (2016) described further aspects of the protest, highlighting the role of moderators (ibid.).

The method applied in the study consisted in scraping the comments made during the conflict, to assign roles to the commenter and to code them for their expression of exit, voice and loyalty, additionally using upvotes as signals for community agreement. Key comments were then used to explain the conflict, using the terms from Hirschman's framework. Various questions were raised, like whether the framework that was developed for a commercial scenario can fully catch all the aspects of an online platform with its additional features and dynamics, and whether it can explain apathy. In any case, Centivany and Glushko (2016) is an interesting example of the application of Hirschman's framework and a demonstration of the importance of policy.

In a study on the warez scene, Chandra describes competition as a collective resource pool that enabled the scene to function and to govern itself for the last 30 years:

> This study argues that with the ludic competition within the warez scene itself an institution for collective action, it can be approached

> as a common-pool resource, which participants use to gain ego boosts or reputation (Chandra, 2016, p. 377).

Using the Ostrom's framework of long-enduring common-pool resource institutions (Ostrom et al., 1994), it describes how that resource can lead to long stretches of participation:

> The study finds that not only can competitive play sustain a community built around it, but as a CPR, it can, with the right institutions, endure in the face of continuous environmental changes as well as individual rational self-interest (Chandra, 2016, p. 381).

As a consequence for designers, Chandra state that the Ostrom's principles can help to assess whether the necessary infrastructure exists to mirror the functioning of the warez scene, i.e. means to enforce community rules, and to adapt to environmental change.

A current trend in participation research is to reframe participation as engagement (e.g Grinberg et al. (2016)). This might be to also cover indirect participation, like signs that a user has thought about something, or it might be just an adoption of language formerly primarily used in marketing.

Grinberg et al. looked at the effect of posting on Facebook. They observed patterns of general activity increase before and after a new post. They suggest to use that effect of heightened interaction to lead users to participate (ibid.).

A study on the usage of animated GIFs found that the high engagement quality (as defined as a high number of likes and reposts on Tumblr) stems mainly from their usability, including an appreciated lack of functionality resulting in smaller file sizes, allowing fast load times (no sound, comparably bad image quality). But the depicted motive and quality of the animation mattered as well. The study also showed that the cultural usage of these short videos, as remixable placeholders for actual and potential emotional reactions, is something special (Bakhshi et al., 2016).

2.4 Conclusion

The topics of identity and anonymity give a very different image than participation when looking at their related literature.

It is possible to form a pretty complete image of participation, and fundamental questions are answered or sometimes obvious. Whether a user will participate in a group depends on how feasible that participation is, on his motivation, his traits ()as defined as his disposition to do something like that), and the use he might get out of his participation. These factors are easily observable - like the usability of a system, the system factors influencing participation, which is still an ongoing process requiring further work, but at least a clearly defined work. Or, these factors can be explained and modified with several sound theories, like common identity and common bond theory.

Identity and anonymity are a lot more complicated. The literature covering identity alone is vast and going back to hundreds of years ago, with many theories and models, that are totally incompatible. They present esoteric ideas that are impossible to prove empirically, but still they appear as useful to explain specific real world occurrences. Anonymity is similarly complicated, but in a different way. As well as identity - also because anonymity covers the question of identity - anonymity has been discussed in the literature as a psychological, political and ethical question for more than hundred years. But additionally, the specific effects of anonymity were studied in laboratory experiments and observed in the field, and online anonymity was especially studied in a few key studies.

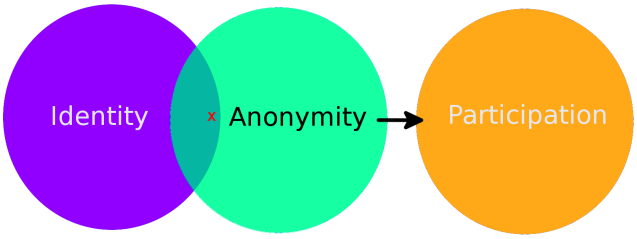

Figure 2.7: Visualizing the topics of this thesis.

The role of this thesis is to look at the link between participation and

anonymity and to examine the effects of anonymity. My analysis of the existing literature led me to identify a missing piece in anonymity research: While there is much theorizing about the effects of anonymity, there are only few studies that have looked at those effects in real world online environments. These studies also conflict: While Kilner and Hoadley (2005) must be understood as a hint to forbid anonymity in future communities, Andalibi et al. (2016) stresses the positive function that anonymity has for people seeking social support. Omernick and Sood (2013), instead, gives conflicting results about the direct impact on participation when disabling anonymity. However, all of this work converge about the pseudonyms: pseudonyms have a very special , as they allow users to be neither fully anonymous nor using their full civil identity, but they can have a similar effect as when they are fully identified (Kilner and Hoadley, 2005). After the general effect of anonymity in real online environments, the role of pseudonyms is the second aspect this thesis will handle. The approach taken for this will be described in the following chapter.

Chapter 3

Approach

3.1 Introduction

So far, chapter 1 on page 6 has explained the research question of how anonymity influences participation in the web. chapter 2 on page 11 identified a gap in the literature about the actual role of identity and pseudonyms, and the lack of studies in existing environments. This motivated my research, and my approach of working in three steps:

1. A literature review
2. Create hypotheses
3. Test them to disprove

This approach was chosen while following a seminar about epistemology, positivism and empirical falsification. The idea was to look at the current state of the knowledge via a literature review, then based on that to generate hypotheses suited to answer the research question, and then to test them in real online environments to only keep those that were not disprovable.

Thus the first step was to do an extensive literature review. This spanned several topics and disciplines and is to a big part presented in chapter 2 on page 11. A second literature review used parts of the first, but was uniquely focused on knowledge about factors influencing participation, and factors influenced by anonymity.

However, in order to relevantly be useful, this knowledge needed to be synthesized. As a second step, I generated a model showing all the factors that influence participation, and which of those factors are influenced by anonymity (see section 3.2). All the links in the model can be understood as hypotheses. Those are not necessarily the explicit hypotheses I wanted to generate, but as hypotheses that could already be tested in the next step.

The third step was to test the model and its contained hypotheses. For that, I looked at existing online environments (see chapter 4 on page 44). Studies of existing online environments are what, according to me, was missing in the literature, especially given the resulting conflicts of the few existing ones as shown in chapter 2 on page 11. The purpose of the studies was two-fold: For one, they allowed to test my model. For two, they also allowed me to generate more complete hypotheses about the influence of anonymity on participation.

The first part of the third step was to look at an environment that recently changed from anonymous to non-anonymous identity models, mirroring studies done before like Kilner and Hoadley (2005) and Omernick and Sood (2013), for which Youtube was chosen. Looking at a changing environment appeared as a good way to see the effect of anonymity, without the possibly differing results of a mixed environment. This study however raised the concern that additional changes influenced participation more than the anonymity change did. Because of that, the second study looked at Quora, an online environment where users can choose their identity model, as they can opt in favor of posting anonymously. Given the interesting results regarding social signals in that second study, the third and final one observed the influence of identity factors on social appreciation, by looking at upvote patterns on Hacker News.

In the following sections of this chapter, we are going to first explain the model that we built and that is the common foundation of our studies. Then, we will present in detail each of the studies we performed.

3.2 Model

In the following section, I first describe the factors that foster participation that we have identified in the literature. The next section highlights the influence of anonymity on interaction, and the final section shows the intersection between the two areas, and the influence of anonymity on participation factors.

3.2.1 Factors that Foster Participation

Anonymity

In Kilner and Hoadley (2005), after progressively changing from anonymous participation possibilities to the need to disclose the full civil identity, they observe that while many metrics measuring participation did not change, what did change was the amount of comments posted. This might be explained by other factors that changed because of that - like a lessened sense of shared identity (see below). However, I believe that it is important not to forget that there is also the possibility of seeing this as a direct influence (see its description in section 2.2.2 on page 16).

Factors from the Common Bond and Common Identity Theory

In Ren, R. Kraut, and Kiesler (2007), the authors argue that attachment to the group influences the participation into the group. They mention two theories explaining that attachment, Common Identity and Common Bond. They write:

> Common identity theory makes predictions about the causes and consequences of peoples attachment to the group as a whole. (ibid., p. 377)

Their literature review highlight several factors that might achieve a group identity, and therefore foster participation:

Social Categorization By just declaring that people are in a common group, based on arbitrary criteria.

Interdependence Being dependent on the other members to achieve a common goal or by a shared fate.

An example for that is described in Ling et al. (2005), where, in an experiment, users contributed more work when they were told that their work is unique and that, thus, it was needed to achieve the group's goal.

Intergroup Comparisons Doing comparisons of members in a group with other groups.

In Ren, R. Kraut, and Kiesler (2007), Common Bond theory is described:

> Common bond theory makes predictions about the causes and consequences of peoples attachment to individual group members. (Ren, R. Kraut, and Kiesler, 2007, p. 377)

Again using literature review, they highlight the following factors:

Social Interaction Direct Interactions between members of a group.

> This fits to Chmiel et al. (2011), where it was later described that the expression of negative emotions lead to high interactions between users and to a high amount of participation in the respective threads on the BBC forum. Also, it fits to the success of personalized invitations, stressing the social aspect of a forum (Harper et al., 2006). This success however was not seen in Sharma et al. (2011); to the contrary, as mentioning social aspects in an invitation lead to less registrations with less filled profiles.
>
> But in Oostendorp and Varik (2011), interaction was rather seen as a metric of a successful community. This shows one of the fundamental problems of this area, which is to distinguish between cause and effect.
>
> Tausczik et al. (2014) describe that depending on the community, direct interaction can be necessary to have a effect, in contrast to only creating social awareness.

Personal Information Opportunities for Self-Disclosure.

> For example in Oostendorp and Varik (2011), the option to have an avatar photo was seen to increase the amount of created messages and forum threads.

Personal Attraction through Similarity Because people like people with whom they share similarities.

The factors of Common Bond and Common Identity seem to profit from design for **Social Presence**, like the integration into teams with visible and shared progress (Farzan et al., 2011):

> The results of our studies support the idea that social presence of others can be manipulated on-screen to foster the formation of common bonds and common identities, and that this visual presence leads to greater commitment to the site and longer and more participation. (ibid., p. 9)

Other Factors

There are many more factors described in the literature that cannot be put under the umbrella of the Common Identity and Common Bond Theory.

Politeness as a factor is described in several places. In Burke and R. Kraut (2008), a small sample of messages from discussion groups was analyzed and their perceived politeness measured by an Internet survey. It was found that polite messages got three times more replies in technical groups, but impolite messages got more replies in political groups. Politeness also goes into the question of civility and the effect discussions have on the functioning of society (Santana, 2012).

Introductions and **Requests** are rhetorical strategies analyzed for their impact on responses in Burke, Joyce, et al. (2007). There, they increased the likelihood of replies by 7% and 6%. But also other rhetorical features like the use of **self-references** were shown to elicit responses (Arguello et al., 2006):

> Posts that included testimonials or requests were more likely to receive a reply. Including self-references (I), third- person pronouns, describing cognitive states and process, and expressing either positive or negative emotions all increased the likelihood that a message received a response. The topical coherence of a message with respect to other recent discussions in the community also affected the likelihood of getting a reply. (ibid., p. 6)

If timely **Feedback** is given and depending on which type of feedback, this increases the effort put into the task at hand or the general motivation (Zhu, Zhang, et al., 2013). Strong negative feedback is shown to decrease motivation to participate (Zhu, R. Kraut, et al., 2012).

The **behaviour of the founder** of a group can influence its chance of success. If he is, for example, too controlling, it was observed that groups die early (R. E. Kraut and Fiore, 2014). Likewise, in the context of a learning community, the amount of prompts in the course material to answer questions by the organizers lead to higher participation of learners (Ahn et al., 2013).

Literature gives special attention to the activation of 'already present' members in a community who do not actively participate (lurkers). Jennifer Preece and Shneiderman (2009) present some factors which might activate those members and the different steps - reading, contributing, and leading - divided into

Table 1 Usability and sociability factors that may influence reading

Usability	Sociability
Interesting and relevant content presented in attractive, well-organized layouts	Encouragement by friends, family, respected authorities, advertising
Frequently updated content with highlighting to encourage return visits	Repeated visibility in online, print, television and other media
Support for newcomers through tutorials, animated demos, FAQs, help, mentors, contacts	Understandable and clear norms or policies
Clear navigation paths so that users have a sense of mastery and control	A sense of belonging based on recognition of familiar people and activities
Universal usability to support novice/expert, small/large display, slow/fast network, multilingual, and users with disabilities	Charismatic leaders with visionary goals
Interface design features to support reading, browsing, searching, and sharing	Safety and privacy

Table 2 Usability and sociability factors that may influence contributing

Usability	Sociability
Low threshold interfaces for easily making small contributions, e.g., no login	Support for legitimate peripheral participation so that readers can gradually edge into contributing
High ceiling interfaces that allow large and frequent contributions	A chance to build their reputation over time while performing satisfying tasks
Visibility for users' contributions and frequency of views; aggregated over time	Recognition for the highest quality and quantity of contributions
Visibility of ratings and comments by community members	Recognition of a person's specific expertise
Tools to undo vandalism, limit malicious users, control pornography and libel	Policies and norms for appropriate contributions

Table 4 Usability and sociability factors that may influence leadership

Usability	Sociability
Leaders are given higher visibility, and their efforts are highlighted, sometimes with historical narratives, special tributes, or rewards	Leadership is valued and given an honored position and expected to meet expectations
Leaders are given special powers, e.g., to promote agendas, expend resources, or limit malicious users	Respect is offered for helping others and dealing with problems
Mentorship efforts are visibly celebrated, e.g., with comments from mentees	Mentors are cultivated and encouraged

Figure 3.1: Factors described in the Reader-to-Leader Framework, taken from Jennifer Preece and Shneiderman (2009)

the categories usability and sociability (see fig. 3.1). In its essence, it follows the thought that **easy access to the means to contribute** and **social appreciation of the contribution** will activate lurkers. On the other hand, users who do not post have a variety of reasons for that, including privacy concerns (Nonnecke and Jenny Preece, 2001), and lurking sometimes can be simply regarded as a metrics showing that the community does not fit to the non-participant (Jenny Preece et al., 2004). Nonetheless, they are sometimes regarded as a strong negative factor for the survival of an online community, for example in Shiue et al. (2010), where it is also proposed that **perceived risks** and **social ties** are sufficient to explain lurking behavior.

3.2.2 Influence of Anonymity on Online Interaction

Anonymity is associated with the ability to change situations, to have an influence on various factors.

A survey of 44 people on the Internet with various backgrounds focused on the self-perceived merits of being anonymous. One of the factors is the **emotional benefit**, allowing them to be more relaxed. Also, they perceived anonymity as enabling them to give more honest ratings or recommendations (see fig. 3.2) (Kang, Brown, et al., 2013).

Category	Advantages of being anonymous	Advantages of being identified
Social connections	Avoid disliked others Avoid commitment to the community Lower barrier to new relationships Protect others one cares about	Connect to real life friends Have stronger social connections Encourages more participation
Reputation and trust	Give honest rating/ recommendation	Good for reputation building Gain trust from other users
Image building	Have control over personal image Avoid embarrassment /judgment /criticism	Avoid harsh criticism Consistent with self-image
Emotional benefit	Feel relax and comfortable Feel cool and sophisticated	Feel real, integrated Feel closer to people
Express opinion	Feel free to express views	Avoid irresponsible behavior
Privacy	Have more control over personal information disclosure	Look innocent
Security	Protect personal safety Avoid legal repercussion/spam/stalk/lost of property	Hide in the crowd
Ease of use	Saves effort to log in	Easy to remember account

Figure 3.2: Perceived tradeoffs of anonymity, taken from Kang, Brown, et al. (2013)

Credibility accordingly seems to be another factor influenced by anonymity. While there are theories going in both directions - more or less credible - in Rains (2007) it was less credible when the perceived anonymity was observed.

In a group situation, whether or not one should follow **Conformity** seems to be at least minimally affected by the perceived anonymity (Tsikerdekis, 2013).

Uncivility and **Impoliteness** is attributed to anonymity as well (Levmore and Nussbaum, 2010).

In Shiue et al. (2010) it was not only stated that inactivity might danger communities (see above), but also that anonymity will result in **stronger social ties**, thus minimizing lurking behaviour.

On the other hand, anonymity is suggested to lead to more **antisocial be-**

havior, in the context of online games to griefing (Chen et al., 2009). A similar phenomenon was observed as well seen in the already mentioned study by Kilner and Hoadley (2005), where the removal of anonymity options led to **fewer antisocial comments**.

But in Kilner and Hoadley (ibid.) whether the change from anonymity to being indentified had an **effect on participation** was also measured. They found there is less direct participation, but the same amount of logins and page views.

In Postmes et al. (2002), an experiment with two groups tested the difference between groups with depersonalized (anonymous) and identified members. They found:

> Depersonalization was associated with greater attitude differentiation than individuation. (ibid., p. 13)

They also **identified stronger** with their own group. In a similar vein, as a fitting summary (Bodle, 2013) states:

> The attributes of anonymity, including minimal accountability, disinhibition, and deindividuation, can encourage robust political speech, provide safety from reprisal, permit the freedom to speak freely, and create a strong sense of group identity. (ibid., p. 30)

3.2.3 Influence of Anonymity on Participation Factors

At this point, we have a list of factors that foster online participation, and a list of factors that are influenced by anonymity. We can now intersect these factors (see table 3.1 on the next page) to see how anonymity might influence participation.

We saw above that a large part of the literature assumes that anonymity influences politeness (Levmore and Nussbaum, 2010). Politeness appears to influence participation, and Chmiel et al. (2011) showed that impolite comments provoked other comments.

The relationship between anonymity and intergroup comparisons and social interaction is indirect via social presence. Farzan et al. (2011) described that the factors linked to common bond and common identity could both profit from social presence, and Tu (2002) showed that anonymity influences social presence.

It is obvious how anonymity makes less personal information available. Personal Information influences participation in the common bond model (Ren, R. Kraut, and Kiesler, 2007) and as a social signal (Oostendorp and Varik, 2011).

Equally, personal attraction through anonymity is a factor in the common bond model that favors participation (Ren, R. Kraut, and Kiesler, 2007). This attraction can be heightened by anonymity by hiding individual differences (Cress, 2005; Sassenberg and Postmes, 2002).

In the common identity model, there is social categorization which is influenced by anonymity (Postmes et al., 2002). Social categorization influences participation according to Ren, R. Kraut, and Kiesler (2007).

Anonymity can change the perception of contributions and can lead to less social appreciation (Rains, 2007). However, social appreciation and specific types of feedback foster participation (Jennifer Preece and Shneiderman, 2009; Zhu, R. Kraut, et al., 2012).

Related to Common Bond and Identity are social ties. Shiue et al. (2010) mentions them as a factor favoring participation, and Shiue et al. (ibid.) states that anonymity helps to develop social ties.

Shiue et al. (ibid.) also mentions perceived risks as a factor inhibiting participation. Perceived risks are influenced by anonymity according to Kang, Brown, et al. (2013).

Table 3.1: Intersection of factors influenced by anonymity and factors influencing participation.

Factor	According to
Social Presence	Tu (2002)
Personal Information	-
Personal Attr. through Similarity	Sassenberg and Postmes (2002)
Perceived Risks	Kang, Brown, et al. (2013)
Social Ties	Shiue et al. (2010)
Social Appreciation	Rains (2007)
Social Categorization	Postmes et al. (2002)
Politeness	Levmore and Nussbaum (2010)

The resulting model (see fig. 3.3 on page 43) serves as an anchor in the approach to quantify the effect of anonymity on participation in online communities. If the models factors are valid in influencing participation, and anonymity truly influences these factors, one could use that model - or, rather, measure markers of the model's factors - to search for differences in anonymous versus

non-anonymous contributions. If these differences exist, one can assume that there will normally be a difference if anonymous participation is allowed. If there are no differences, then that would mean that anonymity is not significantly influencing participation in communities. This is what has been tested in the studies (next section).

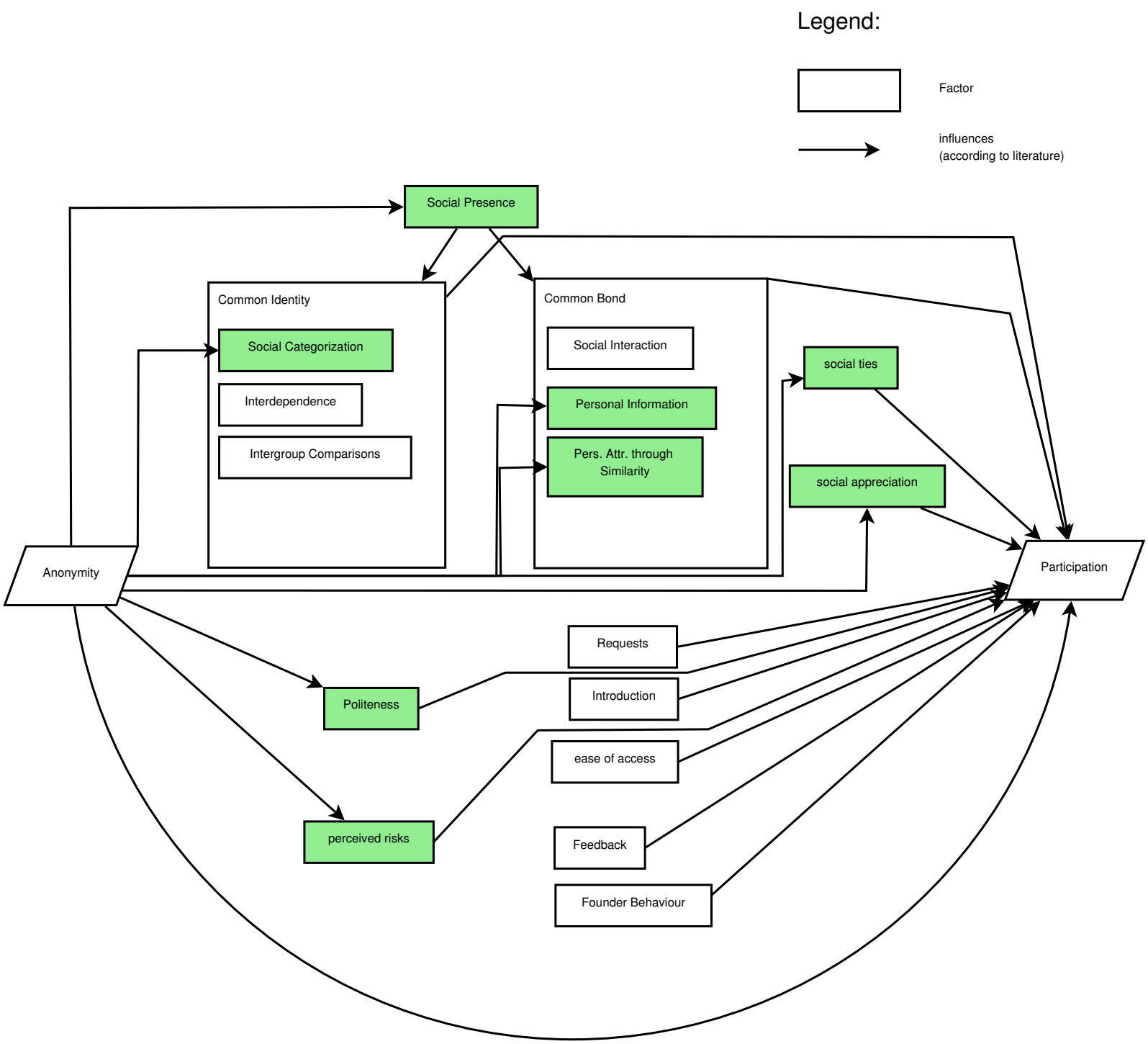

Figure 3.3: The model of factors fostering participation and those influenced by anonymity.

Chapter 4

Studies

With my model at hand, I conducted three studies on existing online communities. The goal was, each time, to measure the connections that the model predicts, but each study had its own focus, progressively developing the research questions. The second study was a reaction to the questions raised by the first study, and the third reacted to the questions of the second.

A common concept that we used in these studies is markers. In fact, studying text-based communication means that we needed a way to detect the presence of each of the factors of the model in text. This is quite possible for some, but seems impossible for others. Take social appreciation as an example: In a platform that has upvotes or other formalized thanks, or even just by closely looking at the written reactions to a post, social appreciation can be measured. But to measure the social ties of a user within a community, solely based on one or a few more comments seems impossible, and the same goes for the perceived risks. As a result, for each of the factors I wanted to measure, I had to find markers that were visible in text.

Which markers existed depended on the investigated platform. This means that, for each study, I created a simplified model with the factors that this specific study could investigate. Those models will be presented with the description of the studies in the following sections.

4.1 Youtube - A changing environment

We decided to look at comments in Youtube because in November 2013, Google integrated Youtube's comment system into Google+. Before the change, users were free to choose a name, but after the change, users were forced to use their full civil identity (later, pseudonyms were allowed, but the character of the platform changed) [1]. Thus, we could find videos with comments made by users with pseudonym only, and newer videos where commenters often used their full name, while being connected by Google+ to their friends and identity.

For Youtube, the phenomena relevant for the simplified model appeared to be politeness, intergroup comparisons and social interaction, with the latter two influenced by anonymity over social presence (see fig. 4.1).

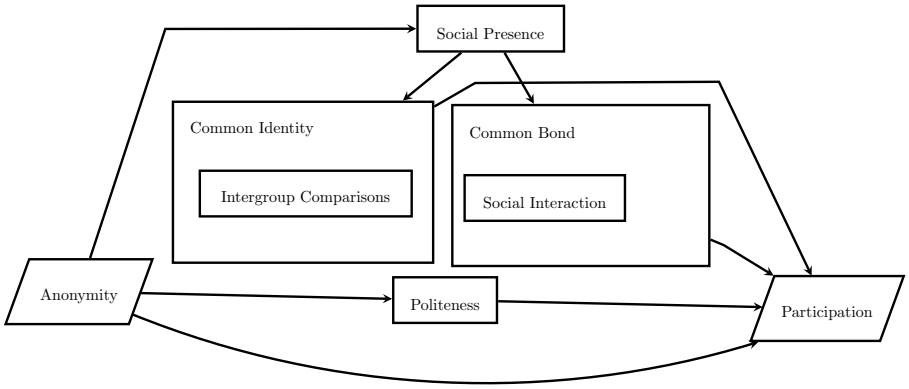

Figure 4.1: The model of participation factors and those influenced by anonymity used in the Youtube study

4.1.1 Data Collection

24 videos were identified. They had several comments and were related to informal caregivers or Alzheimer, topics which we assumed would trigger social support. The average publishing date of the comments was Monday, December 6, 2010. The 3773 comments were downloaded with Youtube's API (using modified scripts of the TubeKit parser[2]), as well as the profile information of the 3087 users, revealing whether the account was linked with Google+ or not. Youtube's

[1] Causing several Youtubers to forbid comments, see http://goo.gl/wkkbBy
[2] http://tubekit.org/

API did not show when users linked their Youtube-Account to Google+; one can only see which commenters are still not using Google+. However, it is complicated to use Youtube while being logged in without going through the Google+ boarding. Consequently, no comment made after the change to Google+ was from a user without Google+. The other way around existed, there were comments from people having only a Google+ account and no Youtube profile, but all were discarded for being formal sharing announcements.

The comments were then analysed for markers that showed:

1. **Politeness**. To measure those factors in text, we examined which markers show how polite a message is. After dismissing some models as too complicated to use manually (House and Kasper, 1981) or not accessible enough (Danescu-Niculescu-Mizil et al., 2013), it was decided to use an algorithmic approach (Bayes' algorithm, see below)

2. **Intergroup Comparisons**. We searched for the words "we/us/our/them", that grammatically show that a group of people is mentioned (Bramley, 2001, p. 86). The hypothesis behind this search is: If people think there are different groups, they will use pronouns to describe these groups. This will happen when they feel being part of one group, one "community", with other groups at the outside of which they are not part, or which are at least distinct from the current group [3]. In the model, the use of intergroup comparisons is influenced by anonymity through **social awareness**.

3. **Personal Interaction**. We looked at the reply count given by Youtube's API. In the model, this is influenced by anonymity through **social awareness**. The amount of replies made has to be fetched from the comment data by searching for the @-character. This metric worked only before the change to Google+.

The use of the Bayes' algorithm was thereby the most complicated step. 300 comments were marked manually by a colleague and me, as either polite, neutral or rude. We empirically defined these categories based on the observation of the corpus of comments, and decided which category was relevant based on our impression of the comment. See below for examples of: a nice comment with thanks and best wishes can be polite, a comment with an insult is impolite, and

[3] Also see http://selp.eu/lexique/pronoms-personnels-2

a comment with no special tone or markers is neutral. To ensure the rigorous categorization of the comments, the analysis was made independently by two reviewers. We kept for the study only those comments where the two reviewers had the same impression. That category was then added to the database. Then the algorithm learnt via a ruby script from our manual classification and classified all remaining comments. The classification of 100 comments was used to calculate an estimated accuracy, manually by one reviewer. The accuracy of the used algorithm was 80%.

Let us now look at this in more detail. To detect the politeness, a script ruby selected 300, printed each out and asked whether they are polite, neutral and rude. One such comment was:

> @songster117 Thank you for your great information!

This comment was marked by both reviewers as polite, since it contains a thanks and praise for another user. Impolite comments were also easy enough to discover, as in this example, which was interpreted as being sarcastic:

> i cried..huhuhu....

The intergroup comparison was detected by another ruby script. In the case of the comment above, it would not have detected as an intergroup comparison, since the comment does not contain the words "we", "us", "our" or "them". But it would have detected as an intergroup comparison in the next comment, where the user is clearly regarding himself as being a part of the group of caregivers:

> What a sad story. What beautiful music. Dementia (Alzheimer's) is a cruel disease for all involved. Caregivers need all the help we can get as we can get physically, emotionally, and mentally ill from the pure stress of it. Thank you for this video! May we find a cure soon!

To detect personal interaction a script searched for @ followed by some characters, meaning that the comment is a direct response to another user:

> @RamjetFilms Where can I see this entire film? Or is this all there is?

4.1.2 Findings

The change

Comparing comments from before and after the Youtube/Google+ change, there is a difference.

Politeness

After the change, one finds slightly more polite and rude comments (see table 4.1), significant by a χ^2-test with $p < 0.01$.

Table 4.1: Change in Politeness

	Polite	Neutral	Rude
Before	133 (3%)	2838 (92%)	155 (5%)
After	32 (5%)	534 (84%)	81 (11%)

Intergroup Comparisons Most of the comments did not contain intergroup comparisons (we/us/our/them). After the change, the average use of those words was slightly higher (see table 4.2, but a t-test showed the increase not to be significant.

Table 4.2: Change of Amount of Comparisons

Group	mean	sd	median	n
Before	0.1628	0.5885	0	3126
After	0.2365	1.3417	0	647

Social Interaction There are two different metrics for social interaction in the data: replies produced replies received. The Youtube's API only shows the amount of replies received. The difference when looking at the effect of the change is important (see table 4.3 on the following page), and significant by t-test with $p < 0.01$. After the change, with an average of 0.5 it seems like every second comment was answered, though the median of 0 shows this to be false. Instead some comments got many replies, while many others still got none.

Table 4.3: Change of replies

Group	mean	sd	median	n
Before	0.0067	0.1171	0	3126
After	0.4791	2.3598	0	647

Pseudonymous vs Google+ Users before the change

The previous section described how the change in the environment had an influence on the comments. But that does not prove that the change in the degree of anonymity is the cause of that change, as other factors changed as well. A difference in the comments between users who adopted Google+ and those who did not, would have been a clearer signal, but the difference was small.

Politeness There was no difference in the politeness rating (see table 4.4), confirmed by a χ^2-test resulting in $p = 0.8424$.

Table 4.4: Politeness of anonymous and Google+ users

	Polite	Neutral	Rude
G+	96 (4%)	2058 (91%)	112 (5%)
pseudonym	36 (4%)	730 (91%)	36 (4%)

Intergroup Comparisons Intergroup Comparisons were also made on the same level (see table 4.5).

Table 4.5: Comparisons of anonymous and Google+ users

	Comments With Comparisons
G+	253 (10%)
pseudonym	83 (10%)

Social Interaction The only visible difference is here. According to the API, no pseudonymous user got any reply (see table 4.6 on the next page). However, they made the same relative amount of replies. The lack of responses could explain why the users stopped being active (Zhu, Zhang, et al., 2013). This observation could be a bug in the API, but is not totally unlikely given the low amount of replies. Comments were often directed at the creator of the video, not at other commenters. Unfortunately, the identification of whether a comment

was a reply or not was not reliable. That data does not come from the API but from searching for an "@" sign, a practice used before the change to reference another user.

Table 4.6: Average Replies of anonymous and Google+ users

	Avg Replies Received	Avg of being a Reply
G+	0.01	0.085
pseudonym	0	0.081

4.1.3 Limitations

It is possible that the markers that were measured are influenced by other factors, and that anonymity did not play a significant role. Youtube changed its interface, the spam control and the ranking of comments, from a timeline system showing the newest comments first to an opaque ranking system. External cultural factors could also influence the comments. Thus a different selection of videos could show other results. Another limitation is the bayes algorithm used to qualify politeness. The initial supervised learning process depends on the researcher entering the input. The observed 80% accuracy is subject to the same limitation, as the algorithmic politeness rating was compared with the subjective correct rating.

4.1.4 Conclusions

The results lead to two hypotheses: (1) When commenters are anonymous, it leads to less polite and less rude comments. (2) When commenters are anonymous, it leads to less interaction.

The first hypothesis is especially surprising, as it stands in contrast to what was found by Kilner and Hoadley (2005). It is further interesting to see that there was no difference observed between the commenters using Google+ now and those who chose to stay pseudonymous, or to abandon Youtube after the change, apart from the reply count. The expectation when looking at that data was to see a difference caused by a different mentality between those accepting Google+ and those who did not.

4.2 Quora - A mixed environment

Having seen in the prior study that anonymity did not have the expected effect, and that its effect was much less important than our hypothesis, the logical next step was to look at an environment where other factors were less able to influence the interaction. The question and answer platform Quora is such an environment, because its user can opt to answer and ask questions anonymously. It enabled us to look at anonymous interaction in the same environment as identified interaction, without the big change of a comment system, or the possible cultural changes which could have occurred on Youtube over the rather large timespan we looked at.

For Quora, like the study on Youtube, the simplified model contained the participation factors: politeness, intergroup comparisons and social interaction, with the latter two influenced by anonymity over social presence. Additionally it contained social appreciation as measured by the upvotes that users can give to answers (see fig. 4.2).

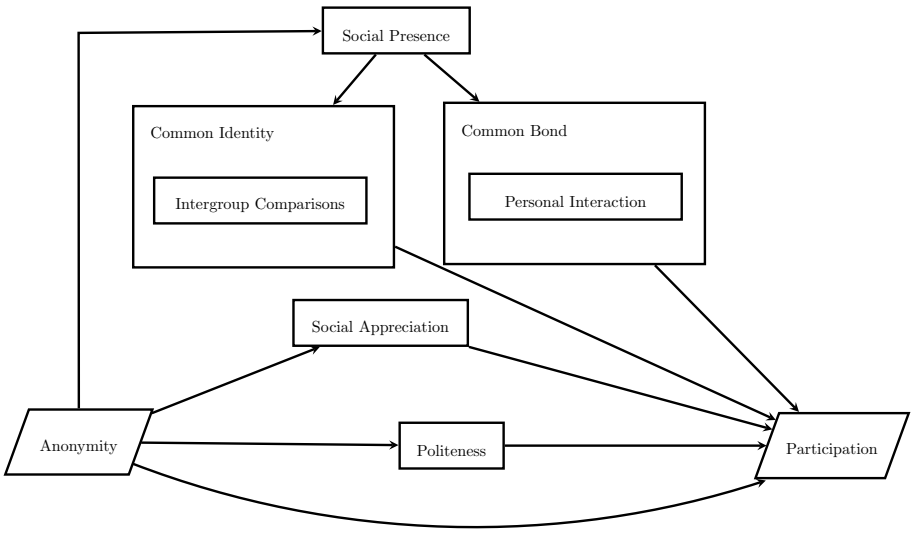

Figure 4.2: The model of participation factors and those influenced by anonymity used in the Quora study

Unlike the prior study, an effort was made to examine, through a survey, users' impression about the anonymity function and how they use it (see section 4.2.1 on page 55). We did that to get additional confirmation of the result

that the anonymous answers seemed equally valued.

This study has also been published in extended form at Paskuda and Lewkowicz (2016), and the previous version in Paskuda and Lewkowicz (2015).

4.2.1 Data Collection and Analysis

We collected data in two distinct steps.

Collection of online data

We selected 375 questions and their 4765 answers, of which 288 were anonymous, by taking the popular questions in the health category (see fig. 4.3 on the following page). Questions without answers or those that merged with multiple other question threads were skipped. The obtained HTML was then parsed, and the generated data were saved in a database. No demographic information about users were collected.

The questions were then categorized by a colleague and I, as belonging to one of the categories *medical*, *lifestyle*, *joke* and *other*. Those categories were chosen after looking at the types of questions in the dataset.

1. Medical questions were defined as relating to a "real" medical health question, like how to react to a cancer diagnosis. An example for this is the question shown in fig. 4.3 on the next page:

   > Depression: What is the diagnosis process like for someone to be diagnosed with depression?

2. Lifestyle questions were the less serious ones like how to stay fit. They constituted the majority of questions. An example:

   > Is sleep overrated? Are 6 hours of sleep enough for a 19-year-old?

3. Joke questions are those with either a clear humoristic intent or asking about a curiosity, like the idea that one could die from taking a cold shower during the day. However, there were not many questions in this category, those questions were thus not analyzed separately. As an example, another shower question:

   > Is it true that you can die if you shower after eating?

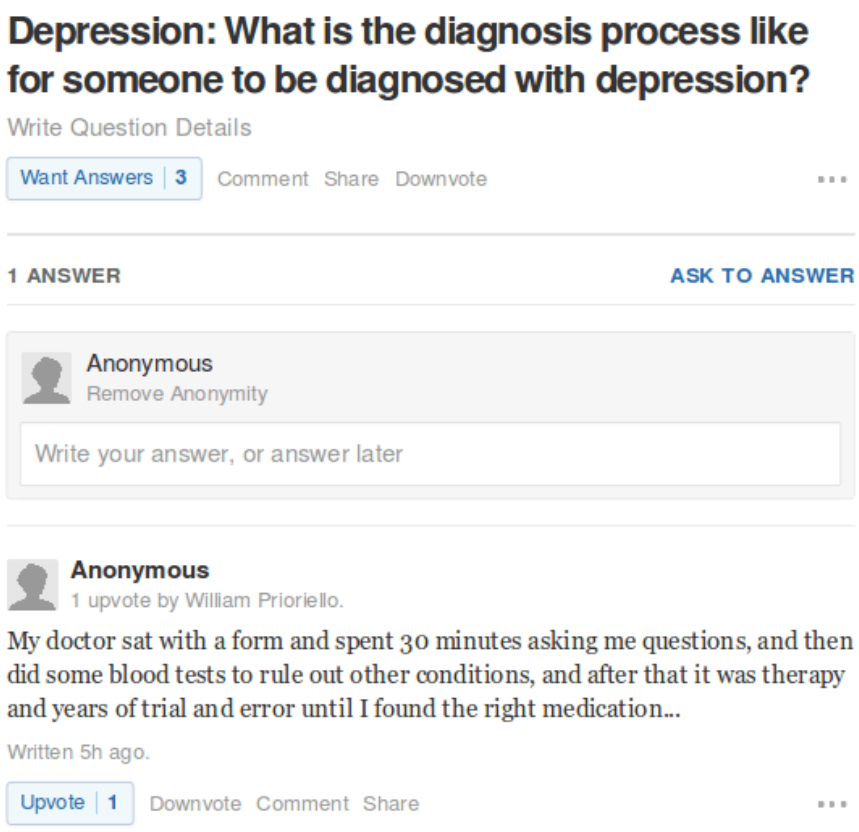

Figure 4.3: Example of an anonymous question and answer on Quora, http://goo.gl/md4WJ3

4. The remaining questions were those with no link to health as such, like:

   What is the best gift for a doctor?

Only the questions where my colleague and I independently agreed upon the category were used in the analysis; 3002 answers remained, of which 148 were anonymous.

The data were analyzed with a number of scripts, in particular, calling a Bayes classifier[4] and a statistic toolkit[5].

We searched for a number of factor markers from our participation model:

1. *Politeness.* How polite the message was. We used a Bayes classifier in an attempt to categorize the answers into the categories *polite*, *neutral* and *rude*. This was based on reports that algorithmic approaches can work acceptably well for detecting politeness (Wild and Stahl, 2007) and our own good experience with the method in the YouTube study.

2. *Intergroup Comparisons.* As in the Youtube study, we searched for the words "we/us/our/them", which show that a group of people is being mentioned (Bramley, 2001, p. 86). To repeat the idea behind this search: If people think there are different groups, they will use pronouns to describe these groups. This will happen when they feel being part of one group, one "community", with other groups at the outside, which they are not part of or which are at least distinct from the current group. In the model, anonymity influences the use of intergroup comparisons through *social awareness.*

3. *Personal Interaction.* To approximate personal interaction, we used the number of comments to an answer. In the model, anonymity influences this through *social awareness.*

4. *Social Appreciation.* The number of upvotes reflected this.

In practice we proceeded as follows:

We categorized all questions and applied the same category to their answers. An example answer from our dataset (see again fig. 4.3 on page 53):

[4] https://github.com/jekyll/classifier-reborn
[5] https://github.com/clbustos/statsample

> My doctor sat with a form and spent 30 minutes asking me questions, and then did some blood tests to rule out other conditions, and after that it was therapy and years of trial and error until I found the right medication...

This answer belongs to what we categorized as a medical question. Depression is a severe illness and as such the categorization was in this case straightforward.

Politeness was seen as neutral. There are no forms of politeness in the answer, no thanks, no specific kind of words or best wishes. There are also no insult in any form, no negative wording, no hidden attack.

Since the answer does not contain "we", "us", "our" or "them", it was not marked as containing intergroup comparisons.

For personal interaction, we looked at the comment counter visible on the page. This answer received no comment, thus this was stored as containing no personal interaction. Of course, given that it is an answer to another user, this decision could be argued against. But in the very least this answer did not foster further direct participation as far as we could see, which interests us the most.

Last is social appreciation. Here, we looked at the upvote counter. The answer got one upvote, which was stored in the database.

The scripts used and the generated database are available at https://goo.gl/GRJ6At.

Survey

In order to get a broader understanding of the use of the anonymity function, we decided to complete the data collected on the platform by an online survey (see table 4.11).

Through 12 questions, this survey asked Quora users to provide their opinion and impression about the anonymity function and its use. This survey was online for several weeks, but feedback only arrived on the first two days, as long as it was linked to a Quora-question.

The survey was answered by seven active Quora users who visit Quora every day of the week. Five of them used the anonymity function. These responses are discussed in section 4.2.2 on page 62 and compared to the prior results.

4.2.2 Findings

We analyzed the answers, which means that we searched for significant differences in the selected markers between anonymously and non-anonymously posted answers, overall and again in their respective category. We used multiple t-tests, which we realized, increases the chance of having made a Type I error. But we did not see many significant differences, and those we saw hold up against error corrections. The result mainly showed that the two groups did not differ greatly, with one noteworthy exception described in section 4.2.2 on page 59.

Politeness

It was not possible for us to algorithmically analyze the answers for politeness as we had done previously. The algorithm failed to distinguish among the three categories (polite, rude, neutral), categorizing nearly all answers as either all rude or all polite while almost ignoring the much more fitting neutral category. This was a surprise given that the same software and workflow were used in the earlier YouTube study, in which we found 80% accuracy.

A manual examination showed that, except for just one answer, all of them followed a specific tone that appeared to be common on Quora. That is not to say that all answers were equal; there was a great range of quality and length in the sample data. Many responses were factual, and others were filled with pathos, but they all lacked easily distinguishable indicators of politeness. These were present in the comments on YouTube, where it was easy to categorize an insult as rude and many best wishes as polite. We assume that this impression is correct and that there was no difference between the groups. We discuss this further in section 4.2.3 on page 63.

Intergroup Comparisons

There was only a small difference in the number of intergroup comparisons made by anonymous and non-anonymous users (fig. 4.4 and table 4.7 on the following page).

The difference was significant with a t-test ($p < 0.05$), which suggests the conclusion that the preference to post anonymously on an otherwise non-anonymous platform does influence identification with the group negatively, at

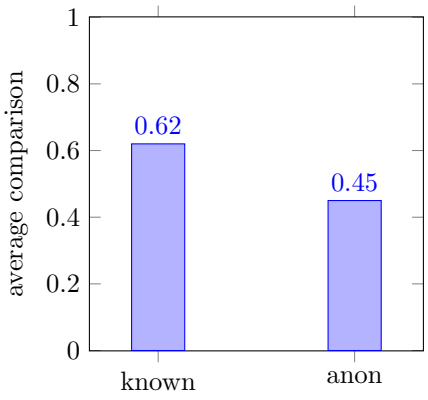

Figure 4.4: Difference comparisons made

least on Quora. The result of $p = 0.01$ also holds up against a Bonferroni Correction of $\alpha = 0.05/4 = 0.0125$, with $m = 4$ for the four initial comparisons without sub-categories. It should however be noted that it would not hold up against a correction with $m = 20$, when including the tests against subcategories.

But the difference was significant only overall, not in the category of medical questions, which is an important category for our research context (informal caregivers sharing experience on the care they provide to their sick relative). The possible hypothesis is then: if the seriousness of the question counteracts the negative effect, then we could ignore the negative effect when designing a system for informal caregivers.

Table 4.7: Amount of comparisons

Group	mean	sd	n	category
Known*	0.62	1.97	4477	all
Anonymous*	0.45	0.98	288	
Known**	0.52	1.84	2854	categorized
Anonymous**	0.28	0.70	148	
Known	0.43	1.39	185	medical
Anonymous	0.36	0.91	25	
Known	0.40	1.34	1550	lifestyle
Anonymous	0.24	0.67	78	
Known	0.53	2.22	927	other
Anonymous	0.33	0.66	40	

Personal Interaction

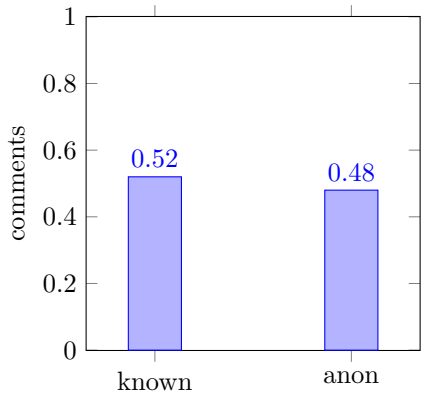

Figure 4.5: Average amount of received comments

A t-test showed no significant difference in the number of comments received for the two groups (fig. 4.5 and table 4.8). In general, comments to answers are not overly common on the platform. In fact, Quora uses elements of the user interface to not highlight them: they are not visible by default, and they can be made visible only by clicking on a small grey-colored link. As such, a rough average of one comment for every second answer was already unexpectedly high.

Table 4.8: Received comments

Group	mean	sd	n	category
Known	0.52	4.45	4477	all
Anonymous	0.48	1.82	288	
Known	0.44	3.55	2854	categorized
Anonymous	0.53	1.71	148	
Known	0.15	0.43	185	medical
Anonymous	0.24	0.52	25	
Known	0.23	1.43	1550	lifestyle
Anonymous	0.46	1.57	78	
Known	0.89	5.91	927	other
Anonymous	0.9	2.39	40	

We should note here that the amount of personal interaction that occurred through direct messages was invisible to us (see also section 4.2.3 on page 63).

The non significance of the difference between the groups was also unexpected. We observed a high increase in personal interaction in the YouTube

study after the change to less anonymous comments. The fact that there was no difference here suggests another explanation: that the change on YouTube was not caused by the change in anonymity but by the change in the comment User Interface and the link with the Google+ social network.

Social Appreciation

The number of received upvotes did not differ significantly between anonymous and non-anonymous questions. However, anonymous answers also received less feedback (fig. 4.6 and table 4.9 on the current page and on the following page).

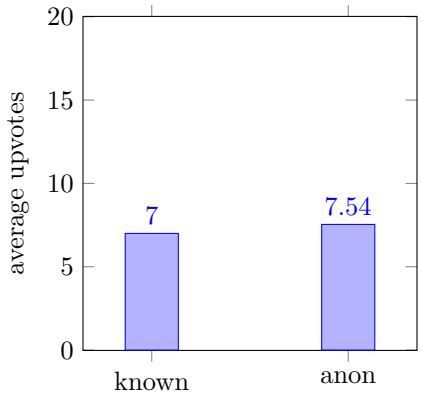

Figure 4.6: Average upvotes received

There was sufficient literature with different results and theories to expect a stronger difference. In particular, Rains (2007) led us to expect that anonymous answers would be less appreciated. In that study, an experiment measured the assigned persuasiveness of responses that were linked to whether a participant was anonymous or not. These anonymous responses were viewed as being inferior to the non-anonymous ones, as less trustworthy and less persuasive. We expected that the same would happen here, i.e., that the anonymous responses would receive fewer upvotes.

However, another factor that was measured was the length of the answer, which by itself was not significant (table 4.10). But the difference between the two groups was significant in a t-test with $p < 0.01$ when looking only at the answers in the lifestyle category. The result of $p = 0.0013$ also holds up against a Bonferroni Correction of $\alpha = 0.05/20 = 0.0025$, with $m = 20$ for all significant

Table 4.9: Upvotes

Group	mean	sd	n	category
Known	7	41.36	4477	all
Anonymous	7.54	60.14	288	
Known	6.6	41.76	2854	categorized
Anonymous	6.58	42.63	148	
Known	1.57	3.24	185	medical
Anonymous	1.6	1.44	25	
Known	2.68	26.69	1550	lifestyle
Anonymous	3.83	22.08	78	
Known	14.39	60.53	927	other
Anonymous	15.7	75.94	40	

tests made.

Anonymous users wrote significantly shorter answers to this kind of question.

Table 4.10: Answer length

Group	mean	sd	n	category
Known	712.3	1025.6	4477	all
Anonymous	743.23	1280.83	288	
Known	678.5	987.72	2854	categorized
Anonymous	747.74	1559.52	148	
Known	666.65	685.98	185	medical
Anonymous	943.48	1243.59	25	
Known**	743.28	911.38	1550	lifestyle
Anonymous**	521.06	557.41	78	
Known	574.82	1190.51	927	other
Anonymous	1033.03	2656.04	40	

Using a Pearson correlation, we found a positive correlation $r = 0.384$ with 286 degrees of freedom ($p < 0.01$) between the length of the answer and the number of upvotes, but only for anonymous answers (fig. 4.7).

This means that for anonymous answers only, the number of upvotes increased with the length of the answer. This is surprising given that that correlation did not exist for the other answers.

It seems plausible to expect that in general, longer answers will receive more upvotes on Quora. They take longer to write, they can contain more relevant information, and they show that an effort was made. The fact that there was no correlation between an answer's length and its number of upvotes shows that

this is not the case. Perhaps longer answers are more cumbersome to read, or perhaps short answers better convey the needed information to answer a typical Quora question.

However, as soon as answers were anonymous, the correlation became significant: Longer answers by anonymous users received significantly more upvotes. Why is that?

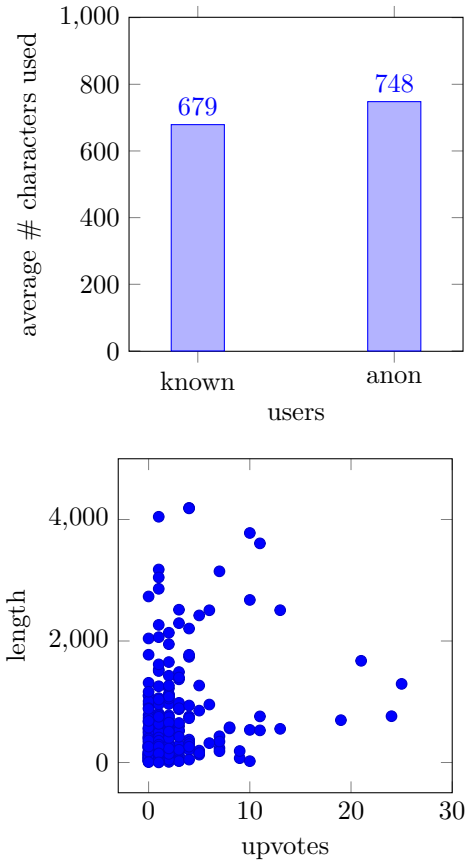

Figure 4.7: Correlations of answers' length and upvotes for anonymous users

A possible explanation is that without the added social cues provided by username and avatar image, readers resorted to comment length as a signal of comment quality.

We base that hypothetical explanation on the assumptions of the reduced social cues approach, as exemplified in Wise et al. (2006). A social signal used

in non-digital communication can, in that model, be replaced by a digital signal. Here, it could be the social dimension - provided by username, attributability and avatar image - that normally influences upvote behavior more than the answer's length. When those signals are missing, a number of things could happen. It could be that the answer's quality becomes more relevant and that quality is linked to length, or it could be that comment length works as a social signal for an answers credibility, a function that was previously covered by the now hidden peripheral elements.

Other explanations are possible. One could transform the removal of social cues into a positive and argue that, as in Cress (2005), removing the avatar image leads people to identify more strongly with those anonymous members. Comment length would again be a more objective factor of comment quality, a factor that was previously covered by negative identification through identity elements such as username and avatar image.

Survey

The responses of the online survey give an interesting view of the impression of the anonymity function by its users. One user said:

> I go anonymous when I'm revealing something that my family members wouldn't want other people to know about

This was the same feeling of the other users, except one who said:

> Out of all the answers I've written on Quora, I can count on one hand the number of answers I've written anonymously. I stand behind the words I write and do not hide behind the mask of anonymity. On rare occasion, I find that there are exceptions. From my point of view, there is rarely a need for it.

Given those statements, the anonymity function is used exactly for the obvious reason: To distance their civil identity from what is said. No secondary function like not getting notifications for example was described.

But how do those users see the quality of anonymous answers?

On average (see table 4.11 on the next page), the quality of the answers on Quora on a Likert-Scale (from 1 to 6) is rated as 4.5. The anonymous answers were rated as 3.7, clearly worse. They are also seen as impolite and half of the survey users do not appreciate getting anonymous answers.

Table 4.11: Survey result

Questions	∅ Answer
Quora visits per week	7
Average Participation	3.9 (max: 6)
Followers	4110
Usage of Anonymity	5 (max: 7)
Own anonymous answers are	scale: 1 to 6
polite	5.4
get comments	3.2
help community identification	2.2
get upvotes	3.4
allow answering some questions	5.2
General answer quality	4.5
Anonymous answer quality	3.7
I appreciate anonymous answers	4.1
I upvote anonymous answers	5.3
Anonymous answers are generally polite	3.1
Happy about getting anonymous answers	3.7

This contrasts with the self-judgements of their own anonymous answers, which they rated to be very polite.

The statement *If I could not answer anonymously, I would not answer some questions at all* got a strong approval of 5.2.

Altogether, it can be said that while users see the anonymity function as useful, they see other anonymous answers as critical. Normal answers are preferred.

4.2.3 Limitations

Quora did not give us raw sample data, and we did not have access to an API. We collected our data manually and then parsed it with a handwritten parser. Thus, the answers that we collected had already been filtered by Quora's moderation, with potential consequences for the validity of the question selection. If, for example, anonymous answers were in general less polite, that effect could have been invisible to us if the rude comments had already been deleted.

Moreover, all of the information that is not available to the public, such as direct messages between members, was also invisible to us. Having this as metadata would have allowed us to measure personal interaction more accurately, and without those data, one should judge the personal interaction metric

as an estimate.

Answers that are submitted anonymously are anonymous only to other users. Quora itself could know who made which answer because users must always be logged in to answer. The answers are, as such, not fully anonymous in the strictest sense because that would include anonymity to all possible observers.

There were not enough anonymous answers in the category "joke" to interpret the results in that category.

Finally, the survey was not filled out by a sufficient number of users for its results to be totally reliable.

4.2.4 Conclusion

Quora provided us with the option to study a mixed environment of anonymous and non-anonymous user-generated texts. In contrast to the YouTube study we conducted earlier, there were fewer differences. On YouTube, the prominent change was on the level of politeness and the increase in social interaction. On Quora, there was only the difference in the correlation between answer length and upvotes for anonymous answers, and a hint for less identification with the community.

Previous work showed the following:

> We know that people who have limited motivation to process content are more likely to base evaluations on peripheral cues (Wise et al., 2006, p. 33)

Something similar could have happened here: Quorans could have normally based their upvotes at least partly on the peripheral social cues provided by username and avatar and resorted to answer length as a relevant factor only when those social cues were not present.

Regarding the effect of anonymity on an online community, I understand the result as an argument for the harmlessness of anonymity. Anonymous answers were, in the eyes of the community in general, not worse, and they did not receive significantly fewer upvotes. Overall, they were not significantly shorter, which could be important for community builders. Moreover, in contrast to the expectations generated by Kilner and Hoadley (2005) and deindividuation theory in general, they were not less polite.

Comparing the results by looking at the different topics in the health category uncovered that the seriousness of the question made almost no difference on the perception of anonymous users. But it showed that anonymous users tend to write shorter answers to non-serious questions.

4.3 Hacker News - A closer look at identity factors

The surprising result of the Quora study was the regression between comment's length and upvotes for anonymous users. This possible influence of social signals convinced me to look closer at identity factors. For that, we chose Hacker News. Hacker News is a US-american tech news site with a strong community aspect: users submit the stories and upvote them, and stories can be commented and these comments upvoted and downvoted. The site does not have an anonymity function like Quora did, but it is also a mixed environment in that some users post under their civil name, while others use a pseudonym not linked to their civil identity.

The goal of the Hacker News study was to further look at the role of identity factors and social signals. With the upvotes for comments at hand, I tried so see whether the social appreciation of a comment on that platform is, in a meaningful way, influenced by the identity factors of its author. In this way, the study did not focus directly on the model, but only on the link between anonymity - if defined as degrees of identity - and social appreciation.

4.3.1 Data Overview

Hacker News has a public API available that makes all public information easily accessible. The administrator of the community also provided us the upvote score (normally hidden) of 50000 comments[6]. The first comment was made on Dec 25, 17:57:36 2015 UTC, the last on Jan 06 20:43:51 2016 UTC. Combined, that gives us:

- For each comment made during these 3 weeks: Its content, author, date, and upvote score.

[6] Given under the promise to not share that data, see section 4.3.3

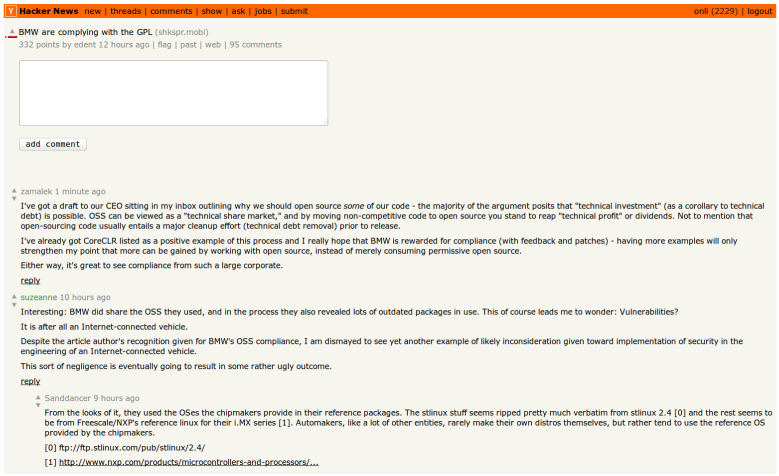

Figure 4.8: Example of a submission with comments on Hacker News, https://goo.gl/9AD1tV

- The profile of each author of a comment, including: When the profile was created, the collected upvotes of all comments and submissions made by this author, and his/her self description (freetext).

We ran a manual outlier detection over the collected data and ran the analysis two times. Since the results did not vary, in the following the dataset without outlier detection will be described.

We decided that the upvote score of each individual comment is our primary output variable, the depending factor. In fact, upvotes are a clear and easy way to identify a signal showing social appreciation, and we used them before (see section 4.2 on page 51 and section 4.1 on page 45). It is also in line with how it is described on Hacker News itself:

> Users should vote for a story because it's intellectually interesting, ...[7]

The second and third depending factors are the amount of submissions and the amount of comments made by a user. They will help us look at the influence of online identity on participation, in coherence with the focus in prior studies we quote (Andalibi et al., 2016; Kilner and Hoadley, 2005; Omernick and Sood, 2013).

[7] as stated in the FAQ

The independent factors are 9 identity factors that signal that the identity disclosed on Hacker News is a real stable identity. These 9 factors do not stem directly from theory or literature; by looking at this specific platform, we identified that they govern how much a user is anonymous or not. Our working assumption is that users who are anonymous will be so according to these factors, and if there is a behavior change caused by anonymity, it will be linked to them.

We have identified that the identity used on Hacker News is close to the civil identity if the user...

1. ...uses a pseudonym or a reference to his civil identity as username.
2. ...mentions an email address,
3. and that email address contains the username.
4. ...mentions a website,
5. and the URL of that website contains the username.
6. ...mentions a Twitter profile,
7. and that twitter profile contains the username.
8. ...has a description in the profile.
9. ...has a profile that existed for many days at the time the comment was made.

Factors 2-8 are detectable in the self description text[8]. Factor 1 is more complicated. Whether a username is a pseudonym or not is not decidable without asking the user. However, what is decidable is whether a username looks like a pseudonym or not, and ultimately this is what we did.

We collected a repository of the most common names used in the last 100 years in the US[9] and picked from 500 randomly selected usernames from our corpus of Hacker News-profiles those that looked like pseudonyms to us. We transformed the names in those repositories into tri-grams with which we learned a bayes-classifier. This classifier then decided whether the name was a pseudonym

[8] The scripts used to detect them can be found, as well as the scripts doing the statistics, under https://goo.gl/LDUT2m

[9] https://www.ssa.gov/oact/babynames/decades/century.html

or a civil identity, and gave us a probability. This was transformed into an identity score, mapping the probabilities for a pseudonym from 0 to 50 and for a civil identity from 50 to 100 (note that x is always negative):

$$f(x) = \begin{cases} x * -0.5 & \text{when pseudonym} \\ \dfrac{x + 100}{2} + 50 & \text{when civil identity} \end{cases} \tag{4.1}$$

This was done for the remaining 12k usernames (see fig. 4.9).

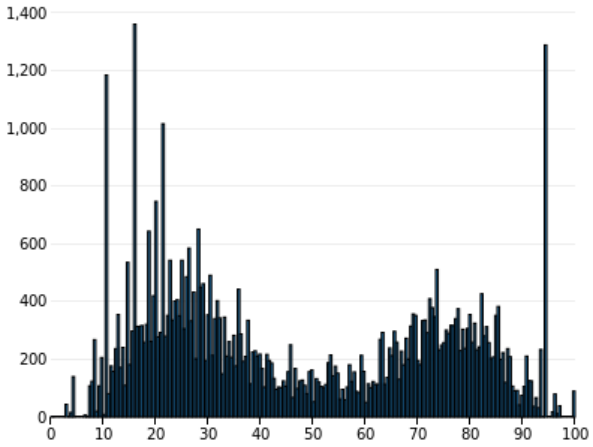

Figure 4.9: Histogram of identity score.

We made a multivariate regression analysis of the binary variables (F2-F8) plus the days of usage (F9) and the rating of the bayes classifier (F1). We also included the comment's length as a dependent variable, because 1) as we reported in the related work section, this length had an important effect in the prior Quora study (see section 4.2 on page 51), where we assumed it served as a social signal and because 2) we know that the algorithm that is placing comments in the threads on Hacker News considers the comment's length as a positive factor, and comments that are placed higher would receive more upvotes. The results are discussed in the following section.

4.3.2 Results

For the three depending variables (upvotes, number of comments and number of submissions), we found a significant regression ($p < 0.001$) within our regression

model. However, the influence of each independent variable varied a lot.

Influence of Identity Disclosure on Upvotes

Upvotes are not influenced a lot by identity factors (see table 4.12). Our model accounted for merely 2% of the variance. But the model showed a significant regression with $F(10, 49960) = 99.711, p < 0.001$.

Table 4.12: Anova summary of upvote correlation

R	R^2	$R^2 Adj$	Std.Error R
0.140	0.020	0.019	6.459

This low correlation prompted us to look in detail at the role each single factor plays in the regression, by looking at their beta coefficients and the individual t-Values. When we look at the effect of the identity factors (see table 4.13), we see that they account to even less of the variance. This follows from their low beta coefficient, which is especially bad for categorical variables. The factor explaining the most is comment length. Comment length seemed to act as a social signal in Paskuda and Lewkowicz, 2015, but it is not an identity factor. It is followed by the number of days since registration. Of the identity factors, we see by looking at the t-value that only the link to a webpage has a significant (positive) effect, and when that link contains the username. But this effect is small.

Table 4.13: Beta coefficients of upvote correlation

coeff	b	beta	se	t
Constant	2.378	-	0.075	31.581
commentLength	0.002	0.131	0.000	29.518
daysOnHN	0.000	0.030	0.000	6.397
hasDescription	-0.059	-0.002	0.120	-0.493
hasEmail	-0.170	-0.009	0.118	-1.446
hasTwitter	0.224	0.009	0.119	1.872
identityScore	0.001	0.002	0.001	0.552
linksWebpage	0.252	0.018	0.079	3.204
mailUsername	-0.223	-0.006	0.191	-1.168
twitterUsername	-1.167	-0.005	1.107	-1.054
webpageUsername	0.443	0.013	0.159	2.788

Link between Identity Disclosure and Number of Submissions

The number of submissions posted on Hacker News correlates stronger (see table 4.14) with our identity factors, $F(10, 49960) = 546.092, p < 0.001$), and its variance is explained to 10%.

Table 4.14: Anova summary of submission correlation

R	R^2	$R^2 Adj$	Std.Error R
0.328	0.107	0.107	267.179

Here, all but one individual factor are significant (see table 4.15). Comment length and the days since registration are, again, important factors. But the identity factors are important as well. The presence of a link to a twitter account has a strong positive effect. What should also be highlighted is the negative effect of the link to a webpage. The only insignificant factor is the upvote score, which was expected, since we observed a difference in user types between people who participate in the platform via comments and via article submissions.

Table 4.15: Beta coefficients of submission correlation

coeff	b	beta	se	t
Constant	-11.861	-	3.145	-3.771
commentLength	0.015	0.027	0.002	6.349
daysOnHN	0.050	0.153	0.001	33.628
hasDescription	110.413	0.097	4.981	22.167
hasEmail	102.112	0.118	4.867	20.982
hasTwitter	194.507	0.181	4.939	39.379
identityScore	-0.382	-0.036	0.045	-8.503
linksWebpage	-13.650	-0.023	3.259	-4.188
mailUsername	-133.175	-0.084	7.896	-16.866
twitterUsername	-145.964	-0.014	45.790	-3.188
upvotes	-0.090	-0.002	0.185	-0.487
webpageUsername	-25.506	-0.018	6.576	-3.879

Link between Identity Disclosure and Number of Comments

The amount of comments posted on Hacker News correlates even stronger (see table 4.16 on the following page) with our regression model, with $F(10, 49960) = 1853.092, p < 0.001$. Its variance is explained to 29% by our factors.

Table 4.16: Anova summary of comment correlation

R	R^2	$R^2 Adj$	Std.Error R
0.538	0.290	0.290	2420.671

As a surprise, the comment length has no strong effect and is also not significant (see table 4.17). But we see all identity factors being used, with overall higher t-Values than in the case of submissions. Having a description, an email and a twitter address correlate with posting more comments, same goes for a name that resembles a civil identity. Linking to a webpage and reusing the username in a link, twitter account, and email address correlate with posting less comments. The upvotes of the comments in our data section had again no effect.

Table 4.17: Beta coefficients of comment correlation

coeff	b	beta	se	t
Constant	-269.261	-	28.497	-9.449
commentLength	-0.010	-0.002	0.021	-0.483
daysOnHN	0.868	0.259	0.014	63.955
hasDescription	1328.773	0.115	45.127	29.445
hasEmail	3483.007	0.396	44.092	78.995
hasTwitter	1510.507	0.138	44.752	33.753
identityScore	2.141	0.020	0.407	5.254
linksWebpage	-341.841	-0.057	29.530	-11.576
mailUsername	-2685.187	-0.168	71.538	-37.535
twitterUsername	-1810.965	-0.017	414.860	-4.365
upvotes	-0.083	-0.000	1.677	-0.050
webpageUsername	-454.102	-0.031	59.577	-7.622

In the next section, we discuss these results regarding the influence of identity and anonymity on participation and social appreciation.

4.3.3 Discussion

Impact of Identity Disclosure on Social Appreciation and Participation

The small effect of the 9 identity factors on the comments' upvotes is not completely surprising. Hacker News is not a site that shows these factors clearly. But it would have been possible that those factors have an effect on how the user writes comments and which comments s/he writes. But that does not seem

to be the case. This also goes for the effect of pseudonymity: Using a civil name or not did not change how comments were valued. But the length of the comment has an effect, which fits to the self-description of Hacker News as a place where commenting should support serious discussions. The fact that the duration since registration is the other significant effect is explainable in a similar way: Users who do not like the discussions on the platform would not comment anymore after a long time, and users who are accustomed to the site culture will be able to phrase their comment in a way that is valued by the community.

Identity factors correlate far stronger with the degree of participation. A user who comments much on the site and provides many submissions is also likely to have a description in his profile (see fig. 4.10), mentioning a email address and a twitter account.

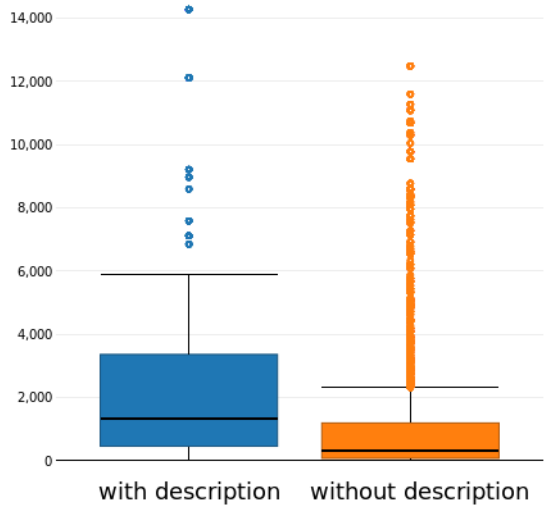

Figure 4.10: Amount of comments of users with and without self-description.

This effect on participation does not necessarily fit prior studies. In Omernick and Sood (2013), the move from anonymous comments to comments under a civil identity via Facebook saw a change in participation, but not an increase. In Kilner and Hoadley (2005), direct participation via posts decreased as anonymous access was reduced. Observing that social appreciation was influenced by identity factors only in a very small way, conflicts with the expectations of SIDE and Deindividuation theory, and with the result of Rains (2007).

The Role of Identity

It is interesting to see that the use of the civil identity as username correlates with less submissions, but with more comments. We have no explanation for that. But the non significance of that factor is surprising and a relevant result. There is a lot of literature which adds to the general opinion that anonymity on the Internet leads automatically to worse discussions, that are less relevant to the topic at hand, less polite and in general of less quality. If that were true, we should have observed a reduced number of upvotes for comments by users with a higher degree of anonymity. Additionally, SIDE and Deindividuation theory state both that anonymous members will identify less or more with a community, which should have been more clearly visible in the regression analysis of the amount of submissions and comments. This is another data point in our research, showing that the negative effects of anonymity, as seen for example in Lapidot-Lefler and Barak, 2012, cannot be clearly confirmed in online settings like Hacker News.

While the other identity factors like a self-description were positively correlated with participation, the link to a webpage and the usage of the username in email address, twitter account and webpage link correlate with less comments and submissions. A possible explanation is the presence of a strong identity on other places on the web. A blogger with his/her own strong identity might choose to write about ideas and publish links on his own blog instead of on Hacker News, and a twitter user might tweet them instead. But without that strong Internet presence, providing personal information could be just a sign of a strong interest for the community. We did not see this tipping point described in the literature before.

Limitations

Our multivariate analysis had a mixture of binary and continuous variables as independent factors, with the dependent variables being continuous. That means that our usage of Anova, while common, violates the assumption of equal variance and normal distribution, lowering the confidence in our results.

We also cannot share the data used here fully, as we agreed with the administrator of Hacker News not to share the comment upvote data, because having it publicly available proved to be toxic to the community before.

4.3.4 Conclusion

Our goal for this study on Hacker News was to quantify the effect of a list of identity factors on social appreciation and participation, following previous work that predicted that there would be an effect, including our own study of Youtube (see section 4.1 on page 45). In this study, we ran a multivariate analysis on 50k comments fetched from Hacker News. We saw significant correlations between identity factors and the received upvotes per comment, and also between identity factors and the amount of comments and submissions made.

The existing correlations that we have identified support the prediction that identity factors impact social appreciation and participation. But that support is not without limitations. Identity factors explained only a small part of the correlation. In particular, on Hacker News, the social appreciation of a comment was not strongly related to the identity factors we selected, which might show a focus on the content of the comment - or just that the factors have to be more visible to directly influence other users.

Our study then shows that pseudonymity, together with anonymity, had a lower influence than predicted and discussed in some of the existing literature Kilner and Hoadley, 2005; Omernick and Sood, 2013; Rains, 2007. We understand our results as further confirmation that the expectation of a big negative influence of anonymity on participation factors is overblown. How users behave depends on too many different factors, and many of those seem to influence users a lot more than their chosen identity.

These results can also be read as a confirmation of SIDE theory in its interpretation that anonymity does lead to behavior that conforms to group expectations, contrary to the assumption that it leads automatically to disinhibition and aggressive behavior. In a community as Hacker News, conforming to group expectations would not necessarily lead to negative behavior.

What was visible however was a correlation between identity factors and participation; an investment in the platform by showing more personal information correlates with more participation, which supports several other results Lampe et al., 2007; Ren, R. Kraut, Kiesler, and Resnick, 2012. It must be noted though that the approach we have adopted does not let us decide which influences what - does higher participation lead to users entering their information, or does entering the information lead later to higher participation?

The additional observation that a strong existing web presence limits par-

ticipation is, to our knowledge, novel, albeit it is not a result that can easily help when designing web platforms.

4.4 Conclusion

After having described our three studies, I can sum up their results as follows:

1. The study on Youtube showed a result conflicting with literature expectations, in that anonymous comments were more polite (and also more often rude, in general more extreme) than those made under Google+ and thus a civil identity.

2. On Quora, I saw no direct difference between anonymous and non-anonymous comments, but interpreted the effect of comment length for anonymous comments as a sign that social signals play a role.

3. The final study on Hacker News investigated further this social signal and focused on the effect of identity factors. The study showed that identity factors influence social appreciation and participation positively, but again, anonymity alone had no significant effect.

Discussion

Taken together, these results conflict with some but confirm some other positions found in the literature.

Anonymity did not result in a breakdown of communications, in incivility or impoliteness. I will make a distinction here between pseudonymity and anonymity. On Youtube, the comment system before the move to Google+ has to be regarded as a pseudonymous identity system. In Kilner and Hoadley (2005), the move from pseudonyms to a civil identity had no big effect. And that is something we saw on Youtube again, that the change from pseudonyms to a civil identity had no clear positive or negative effect.

However, on Quora we saw that anonymous (and not pseudonymous) comments were not more negative neither. In the politeness and social appreciation factors, and while those upvotes can be seen as a proxy for comment quality, the anonymous comments where on par with the normal ones. And on Quora, the normal user accounts are at least pseudonymous; users, often, even seem to

use their civil identity. There also, the negative effect of anonymity could not be seen.

A bit different are the results of the study on Hacker News. But also there, social appreciation did not differ purely on the identity model that has been used. Our computed identity score shows the closeness of a name to a pseudonym or a civil identity.

Thus, our first result is this:

Result 1 *Anonymity does not automatically lead to impolite and uncivil discussion.*

What could be observed on Hacker News was the influence of identity and other factors. User behavior differed, based on what information, if any, a user entered in his profile. Those information also had a measurable influence on social appreciation, measured by the number of upvotes received. Since the profile information are not visible directly next to a comment, I assume that this happens indirectly: the information do not influence other users, they influence the user himself or are a sign of an existing positive attitude towards the platform, resulting in better participation.

This result concurs with the result of the Youtube study. There is sufficient literature to assume that anonymity alone does have a negative influence (Kilner and Hoadley, 2005; Lapidot-Lefler and Barak, 2012; Omernick and Sood, 2013). But since we did not see those negative results in our studies, it is likely that this effect was covered by other factors. For Youtube, that is easily explainable by the additional changes that were done when the system was switched. Not only the identity model changed, also the UI and the sorting algorithm. The situation of being forced into the new system might also affect users negatively. The fact that those other factors can override the possible influence of anonymity is the next result:

Result 2 *Other factors than anonymity have a stronger influence on participation.*

Result 3 *Anonymity can make the effects of social signals visible, e.g. text properties like length influencing social appreciation.*

Therefore, which design implications do these results have for future systems, in particular for the social support platform TOPIC that I presented in the introduction?

Based on this result, we recommend that online software designed for social support should enable an anonymous identity model. It also could use pseudonyms, to minimize potential negative influences caused by anonymity. There are advantages of anonymity, the obvious is not being linked to what was said, that, it is argued, are highly positive in these contexts. Also, the Quora users questioned in the small survey clearly stated the utility of the anonymity function (see section 4.2.2 on page 62). In more general contexts, anonymous discussions could be useful for the general democratic society (Bodle, 2013), especially if the negative effects are less pronounced than suggested in Santana (2012), as observed here.

However, a lot of attention has to be paid to the other factors influencing participation. Because, while we did not observe a strong negative effect of anonymity, we still observed that anonymity and identity factors have the power to influence participation. Together with the prior results of anonymity significantly influencing participation (Andalibi et al., 2016; Kilner and Hoadley, 2005; Lapidot-Lefler and Barak, 2012; Omernick and Sood, 2013), it becomes clear that the other factors are very powerful. If an avatar image can influence communities (Cress, 2005), then designers have to invest a lot of attention in the details that govern influences of social behaviour on the web.

Finally, the impact of the avatar image is linked to the effect social signals had. On Quora, only anonymous comments were appreciated more when they were longer. This uncovering of desirable attributes of submitted texts is something designers can use in various cases. Whenever the focus shall be completely on the text itself, the removal of social signals including the removal of the civil identity and even pseudonyms can be helpful to filter out irrelevant factors.

Additional Results

There are some additional results found during this thesis that do not relate directly to anonymity, but are worth mentioning.

The first one is a confirmation of Lampe et al. (2007). In the study of Hacker News, we observed that users who shared more identity factors in their profile participated more on the platform. Based on this study, it is impossible to determine a causation, but the correlation is a result by itself:

Result 4 *Participation is linked to profile completeness.*

An interesting observation which is, to my knowledge, novel, is the effect of having another home online. In general, users on Hacker News who had a filled in profile were more active and their comments more appreciated. But this did not hold true for users who linked to a webpage and a twitter profile. The observations seems to be as if users with a strong presence elsewhere will be less active on another platform:

Result 5 *An established web presence elsewhere limits participation.*

Chapter 5

Conclusion

In this work, I looked at the effect of anonymity on participation in online communities. The starting point was the design of TOPIC, an online platform for informal caregivers, which I witnessed. During this design, I realized that there is little knowledge about whether users should be anonymous or not. Indeed, the designers of the platform saw clear advantages in this specific context, but were wary of possible negative influences.

Anonymity here means participating under another name than ones' civil identity. It can be a pseudonym or a completely anonymous identifier, like on Quora, where anonymous answers are attributed to *Anonymous*. Participation, as measured here, refers to members of a community writing posts in online discussion boards.

The current state of HCI and CSCW literature is divided on this issue. While Kilner and Hoadley (2005) must be understood as a hint to forbid anonymity in future communities, Andalibi et al. (2016) stresses the positive function that anonymity has for people seeking social support. Omernick and Sood (2013) instead gives conflicting results about the direct impact on participation when disabling anonymity. The existing work already covers the subject of pseudonyms: they have a very special role, as they allow users to be neither fully anonymous nor be compelled to use their full civil identity, but they can have a similar effect as when they are fully identified (Kilner and Hoadley, 2005).

There are two central theories usually used to explain the effect of anonymity:

The first one is classical **Deindividuation theory** (Postmes et al., 2002),

going back to theories about the function of groups from the 19th century (Le Bon, 1896; Reicher et al., 1995). Its main idea is that the individual norms of one person get lost when that person is in a sufficiently large group. Through anonymity and loss of personal responsibility, the single person in the group reverts to primitive and hedonistic behavior, resulting in typical mob behaviour.

The second theory is the **social identity model of deindividuation effect** (SIDE) (Reicher et al., 1995). SIDE can also be seen as a deindividuation theory, but it explains the mechanisms and outcomes of anonymity in groups differently (Cress, 2005). According to SIDE, members of a group do not only lose their social norms, but they adapt to the norms of the group. Those norms can be in conflict with societal norms, but it is not a reversal to a primitive normless state. Also, anonymity works differently in this model: it minimizes the differences between the other individuals, allowing a higher identification with the group. Anonymity strengthening group identification can then equally result in behavior varying from societal norms, but it also favors the group's norms, which are not necessarily negative.

This work provides an insight into the effect of anonymity on participation via a created model and three studies in existing online communities. Chapter 1 on page 6 gave an introduction and explained the research question. Chapter 2 on page 11 was a review of the existing literature on anonymity, identity and participation, concluding with a description of the gap in the literature regarding practical effects of anonymity. Chapter 3 on page 33 presented the chosen approach, and explained and presented the model that was used in the studies. Those studies were shown in the following chapter 4 on page 44, with a detailed explication of the studies on Youtube, Quora and Hacker News. That chapter also contained a discussion of the results when taking all three studies together, but the overall results of the thesis will be presented in this chapter. And just at the end, section 5.4 on page 84 will give a short outlook of possible future work.

5.1 Results

The results of the studies were presented in chapter 4. This section presents them again, and takes the additional step of setting them in the context of the current state of the art in the literature, so as to discuss my contribution

through this thesis.

5.1.1 Anonymity does not automatically lead to impolite and uncivil discussion

The existing literature gave me the expectation that anonymity would have a clear negative effect, noticeable especially by impoliteness, overall worse quality of the discussion and less appreciation for anonymous comments (Kilner and Hoadley, 2005; Omernick and Sood, 2013; Rains, 2007; Santana, 2012). However, in our studies, this hypothesis proved wrong.

On Youtube, Google changed the comment system from one allowing anonymous and pseudonymous user identifiers to one enforcing civil identity via Google+. We did not see a clear negative effect in the comments we analyzed between before and after the switch. What happened was an increase in extreme comments, both polite and impolite. I believe it is likely that other factors, like the change in comment ranking, had a bigger effect than anonymity, and that this comment ranking system preferred comments with more reactions, which might have been the more extreme ones. Therefore, a potential limitation, of which I am aware, is that that those additional changes might affect our results.

Thankfully, the study on Quora did not share that limitation. On Quora, users can choose to post their answers anonymously. There was no additional big change in the platform as there has been on Youtube. The study on Quora showed no negative effect of this feature: Anonymous comments showed no noticeable decrease in quality or politeness, and by looking at their upvotes, they seemed to be equally appreciated.

What we did see was a small effect of several identity factors in a study on Hacker News. On Hacker News, users can participate by writing comments or by submitting links, and they can describe themselves in their profile. We saw a small correlation between the kind of information entered, that we defined as identity factors, and the amount of participation. But we did not see a significant drop in social appreciation of that participation, or otherwise on overall quality.

5.1.2 Other factors than anonymity have a stronger influence on participation

This result goes together with the third result, that *anonymity can make the effect of social signals visible, e.g. text properties like length influencing social appreciation*. What we observed is that anonymity had an effect, by omission. On Quora, the social appreciation of anonymous answers only correlated with their length. On Hacker News, that correlation existed for both, anonymous and non-anonymous comments. Assuming this effect is valid and that length ought to correlate with upvotes, this means that for the non-anonymous comments upvotes, there were other factors that influenced upvotes.

Length could either be a social cue (Wise et al., 2006), or longer comments could just be considered better. Or one could interpret the removal of social cues as something positive and argue that, as in Cress (2005), removing the identity information like the avatar image leads people to identify more strongly with anonymous members. Comment length would again be a more objective factor of comment quality, a factor that was previously covered by negative identification through identity elements such as username and avatar image.

5.2 Implications

If we look at the concept and the theories around anonymity and participation, the result of this thesis clashes with some of them. It stands in contrast to the expectations of the online disinhibition theory:

> However, the disinhibition is not always so salutary. We witness rude language, harsh criticisms, anger, hatred, even threats. Or people visit the dark underworld of the Internetplaces of pornography, crime, and violenceterritory they would never explore in the real world. We may call this toxic disinhibition. (Suler, 2004, p. 321)

This toxic disinhibition might very well exist, but we did not see it as a relevant factor in online communities. There is no doubt that anonymity can be used to attack communities, to create troll-accounts and sabotage discussions, and that it also practically creates a shield from behind which users could say things they otherwise would not say, including insults and personal attacks. Vandalism on Wikipedia for example was observed to come more often from

users without accounts, only showing their IP (Frard et al., 2010). But in the everyday life of communities, these concerns do not seem to be relevant in the actual discussions that take place – though they could be very relevant in communities who lack proper technical defences against these basic types of attacks, like not having means to detect freshly created accounts used to troll.

Regarding the existing studies, these results support Andalibi et al. (2016) and parts of Kilner and Hoadley (2005). For the former, the positive use of anonymous throwaway accounts is something that can also be observed: it is supported by the general positive use of anonymity we observed in those three platforms. For the latter, while we did not see the decline in negativity when going to anonymous to pseudonymous accounts, we did see a minimal non-change with regards to the quality when going from pseudonyms to civil identity. However, the big caveat of Kilner and Hoadley (ibid.) is that the identity is not a civil identity; it is a soldier's militaristic identity, with the army having much stricter rules concerning what one is allowed to say as compared to the rules of the civil society. This could have influenced the discussions greatly from the moment the first link to the militaristic identity was created, even if users were still allowed to use pseudonyms.

I understand the results as a confirmation of SIDE theory. The changes with regards to comments' length when being anonymous or not can be easily explained in the frameworks of that theory, as shown by Cress (2005). The fact that we observed this effect in the online wild through our studies further confirms that explanation model.

5.3 Limitations

There are a number of limitations to this work.

First to mention are the limitations of each study, which are explained in further detail in their respective description in chapter 4. In short, on Youtube it was difficult to measure whether the changes stem from the new identity model or from other changes, on Quora we did not have access to raw data, and the analysis on Hacker News can be attacked on the level of the statistics that have been used(Anova with a mix of binary and continuous factors).

As a whole, this work is based solely on quantitative methods. Further insights might be generated by using qualitative methods to work directly with

the users, to not only observe the real usage but also the inner motivation and reasoning.

This reaches into possible future work, and part of what was planned initially. Future work possibilities will be discussed in the following, final section.

5.4 Future Work

The starting point of this work was the observation that there is a lack of practical recommendations for when and how to use anonymity in online platforms, which was later strengthened when looking at the existing research. Accordingly, this work focused on observing real consequences of anonymity in real existing online environments. At this point, as presented above, I have a number of results about the influence of anonymity on participation.

However, the studies that I realized have the usual limitations that arise when doing studies in the online wild. It is hard to pinpoint that changes are caused directly by anonymity and not by other factors. Part of the process of this thesis was to realize that this complication exists and to react to it: We changed from an environment and situation where it was a big problem that other factors might influence our results (Youtube) to one where the environment was stable, but where we were lacking raw data (Quora) to finally one where we had raw data and the environment was stable (Hacker News). But in that last environment it was still a challenge to properly define the degree of anonymity of users.

Future work should take the result of this thesis and try to define tests to confirm them. I see two ways for that. The first approach would be to go back into the laboratory and to do artificial experiments, but to try to keep them as close as possible to situations that occur online, for instance by emulating real online environments. The second approach would be to do clear-cut experiments in online environments: To generate online communities in which one half is invisible to the other, to let only the first half be anonymous and then to look at the differences in the created communities. That latter approach was my original plan for the end of my thesis – which eventually failed because there were not enough users to get reliable results – for the AAL TOPIC project that was the designated test platform.

Another option is to do inter-platform comparisons. There are several sites

and online forums that have the same topic, but different identity models. An example would be Hacker News, Reddit's technology or programming subreddits, and 4chan's /g/. Future studies could compare communities like this and compare their participation characteristics. It would be necessary to do enough of these studies to, not only measure differences in culture and impact of UI but to see the impact of anonymity. But at the very least, studies like these could be used to further disprove the notion that anonymity has automatically and overwhelmingly a negative effect.

Future work could also focus on the design of User Interfaces, software and communities that embrace anonymity.

Another approach would be to change the method. The method picked for this thesis is mostly quantitative. But it is certainly possible to work with a more qualitative approach, as shown on a very small scale in section 4.2.2 on page 62, or in other publications like Nagel and Frith (2015). Be it surveys or observation, other methods could enrich these results with the users' point of view and deeper reasoning about what happens in specific cases.

Bibliography

Ahn, June, Cindy Weng, and Brian S. Butler (2013). "The Dynamics of Open, Peer-to-Peer Learning: What Factors Influence Participation in the P2P University?" In: *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pp. 3098–3107. DOI: 10.1109/HICSS.2013.515.

Andalibi, Nazanin et al. (2016). "Understanding Social Media Disclosures of Sexual Abuse Through the Lenses of Support Seeking and Anonymity". In: Proceedings of the ACM Conference on Human Factors in Computing Systems.

Arguello, Jaime et al. (2006). "Talk to Me: Foundations for Successful Individual-group Interactions in Online Communities". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06. Montréal, Québec, Canada: ACM, pp. 959–968. ISBN: 1-59593-372-7. DOI: 10.1145/1124772.1124916. URL: http://doi.acm.org/10.1145/1124772.1124916.

Bakhshi, Saeideh et al. (2016). "Fast, Cheap, and Good: Why Animated GIFs Engage Us". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 575–586. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858532. URL: http://doi.acm.org/10.1145/2858036.2858532.

Bargh, John A., Katelyn Y. A. McKenna, and Grainne M. Fitzsimons (2002). "Can You See the Real Me? Activation and Expression of the True Self on the Internet". In: *Journal of Social Issues* 58.1, pp. 33–48. ISSN: 1540-4560. DOI: 10.1111/1540-4560.00247. URL: http://dx.doi.org/10.1111/1540-4560.00247.

Bernstein, Michael S et al. (2011). "4chan and/b: An Analysis of Anonymity and Ephemerality in a Large Online Community". In:

Birnholtz, Jeremy, Nicholas Aaron Ross Merola, and Arindam Paul (2015). ""Is It Weird to Still Be a Virgin": Anonymous, Locally Targeted Questions on Facebook Confession Boards". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, pp. 2613–2622. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702410. URL: http://doi.acm.org/10.1145/2702123.2702410.

Bodle, Robert (2013). "The Ethics of Online Anonymity or Zuckerberg vs. "Moot"". In: *SIGCAS Comput. Soc.* 43.1, pp. 22–35. ISSN: 0095-2737. DOI: 10.1145/2505414.2505417. URL: http://doi.acm.org/10.1145/2505414.2505417.

Borcea-Pfitzmann, Katrin et al. (2006). "What user-controlled identity management should learn from communities". In: *Information Security Technical Report* 11.3, pp. 119–128. ISSN: 1363-4127. DOI: http://dx.doi.org/10.1016/j.istr.2006.03.008. URL: http://www.sciencedirect.com/science/article/pii/S1363412706000343.

Bramley, N.R. (2001). *Pronouns of Politics: The Use of Pronouns in the Construction of 'self' and 'other' in Political Interviews*. Australian National University. URL: http://hdl.handle.net/1885/46225.

Burke, Moira, Elisabeth Joyce, et al. (2007). *Introductions and Requests: Rhetorical Strategies That Elicit Response in Online Communities*. DOI: 10.1007/978-1-84628-905-7_2.

Burke, Moira and Robert Kraut (2008). "Mind your Ps and Qs: the impact of politeness and rudeness in online communities". In: *Conference on Computer Supported Cooperative Work*, pp. 281–284. DOI: 10.1145/1460563.1460609.

Centivany, Alissa and Bobby Glushko (2016). ""Popcorn Tastes Good": Participatory Policymaking and Reddit's". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 1126–1137. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858516. URL: http://doi.acm.org/10.1145/2858036.2858516.

Chandra, Priyank (2016). "Order in the Warez Scene: Explaining an Underground Virtual Community with the CPR Framework". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 372–383. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858341. URL: http://doi.acm.org/10.1145/2858036.2858341.

Chang, Daphne et al. (2016). "Engineering Information Disclosure: Norm Shaping Designs". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 587–597. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858346. URL: http://doi.acm.org/10.1145/2858036.2858346.

Chen, Vivian Hsueh-Hua, Henry Been-Lirn Duh, and Chiew Woon Ng (2009). "Players Who Play to Make Others Cry: The Influence of Anonymity and Immersion". In: *Proceedings of the International Conference on Advances in Computer Enterntainment Technology*. ACE '09. Athens, Greece: ACM, pp. 341–344. ISBN: 978-1-60558-864-3. DOI: 10.1145/1690388.1690454. URL: http://doi.acm.org/10.1145/1690388.1690454.

Chmiel, Anna et al. (2011). "Negative emotions boost user activity at {BBC} forum". In: *Physica A: Statistical Mechanics and its Applications* 390.16, pp. 2936–2944. ISSN: 0378-4371. DOI: http://dx.doi.org/10.1016/j.physa.2011.03.040. URL: http://www.sciencedirect.com/science/article/pii/S0378437111002731.

Cress, Ulrike (2005). "Why Member Portraits Can Undermine Participation". In: *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!* CSCL '05. Taipei, Taiwan: International Society of the Learning Sciences, pp. 86–90. ISBN: 0-8058-5782-6. URL: http://dl.acm.org/citation.cfm?id=1149293.1149305.

Danescu-Niculescu-Mizil, Cristian et al. (2013). "A Computational Approach to Politeness with Application to Social Factors". In: *CoRR* abs/1306.6078. URL: http://arxiv.org/abs/1306.6078.

Davenport, David (2002). "Anonymity on the Internet: Why the Price May Be Too High". In: *Commun. ACM* 45.4, pp. 33–35. ISSN: 0001-0782. DOI: 10.1145/505248.505267. URL: http://doi.acm.org/10.1145/505248.505267.

Derks, Daantje, Agneta H. Fischer, and Arjan E.R. Bos (2008). "The role of emotion in computer-mediated communication: A review". In: *Computers in Human Behavior* 24.3. Instructional Support for Enhancing Students' Information Problem Solving Ability, pp. 766–785. ISSN: 0747-5632. DOI: http://dx.doi.org/10.1016/j.chb.2007.04.004. URL: http://www.sciencedirect.com/science/article/pii/S0747563207000866.

Edman, Matthew and Blent Yener (2009). "On Anonymity in an Electronic Society: A Survey of Anonymous Communication Systems". In: *ACM Comput. Surv.* 42.1, 5:1–5:35. ISSN: 0360-0300. DOI: 10.1145/1592451.1592456. URL: http://doi.acm.org/10.1145/1592451.1592456.

Farzan, Rosta et al. (2011). "Increasing Commitment to Online Communities by Designing for Social Presence". In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. CSCW '11. Hangzhou, China: ACM, pp. 321–330. ISBN: 978-1-4503-0556-3. DOI: 10.1145/1958824.1958874. URL: http://doi.acm.org/10.1145/1958824.1958874.

Frard, Dominique et al. (2010). "The role of argumentation in online epistemic communities: the anatomy of a conflict in Wikipedia". In: *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics*. ACM, pp. 91–98.

Goffman, Erving (1949). "Presentation of Self in Everyday Life". In: *American Journal of Sociology* 55, pp. 6–7.

Grinberg, Nir et al. (2016). "Changes in Engagement Before and After Posting to Facebook". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 564–574. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858501. URL: http://doi.acm.org/10.1145/2858036.2858501.

Harley, David and Geraldine Fitzpatrick (2012). "Appropriation of social networking by older people: two case studies". In:

Harper, F. Maxwell et al. (2006). *Talk Amongst Yourselves: Inviting Users To Participate In Online Conversations*.

Hollenbaugh, Erin E. and Marcia K. Everett (2013). "The Effects of Anonymity on Self-Disclosure in Blogs: An Application of the Online Disinhibition Effect". In: *Journal of Computer-Mediated Communication* 18.3, pp. 283–302. ISSN: 1083-6101. DOI: 10.1111/jcc4.12008. URL: http://dx.doi.org/10.1111/jcc4.12008.

House, J and G Kasper (1981). *Politeness markers in English and German, in Conversational routine: Explorations in Standardized communication patterns and prepatterned*.

Joinson, Adam N (2001). "Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity". In: *European journal of social psychology* 31.2, pp. 177–192.

Jung, Carl Gustav (1953). *Collected works of CG Jung: Vol. 7; two essays on analytical psychology*. Routledge & Kegan Paul.

Kang, Ruogu, Stephanie Brown, and Sara Kiesler (2013). "Why Do People Seek Anonymity on the Internet?: Informing Policy and Design". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, pp. 2657–2666. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481368. URL: http://doi.acm.org/10.1145/2470654.2481368.

Kang, Ruogu, Laura Dabbish, and Katherine Sutton (2016). "Strangers on Your Phone: Why People Use Anonymous Communication Applications". In: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. CSCW '16. San Francisco, California, USA: ACM, pp. 359–370. ISBN: 978-1-4503-3592-8. DOI: 10.1145/2818048.2820081. URL: http://doi.acm.org/10.1145/2818048.2820081.

Kiene, Charles, Andrs Monroy-Hernndez, and Benjamin Mako Hill (2016). "Surviving an "Eternal September": How an Online Community Managed a Surge of Newcomers". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 1152–1156. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858356. URL: http://doi.acm.org/10.1145/2858036.2858356.

Kilner, Peter G. and Christopher M. Hoadley (2005). "Anonymity Options and Professional Participation in an Online Community of Practice". In: *Proceedings of th 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!* CSCL '05. Taipei, Taiwan: International Society of the Learning Sciences, pp. 272–280. ISBN: 0-8058-5782-6. URL: http://dl.acm.org/hcitation.cfm?id=1149293.1149328.

Ko, Minsam et al. (2016). "Lock N' LoL: Group-based Limiting Assistance App to Mitigate Smartphone Distractions in Group Activities". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 998–1010. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858568. URL: http://doi.acm.org/10.1145/2858036.2858568.

Kraut, Robert E. and Andrew T. Fiore (2014). "The role of founders in building online groups". In: *CSCW*, pp. 722–732.

Lampe, Cliff A.C., Nicole Ellison, and Charles Steinfield (2007). "A Familiar Face(Book): Profile Elements As Signals in an Online Social Network". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: ACM, pp. 435–444. ISBN: 978-1-59593-593-9. DOI: 10.1145/1240624.1240695. URL: http://doi.acm.org/10.1145/1240624.1240695.

Lapidot-Lefler, Noam and Azy Barak (2012). "Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition". In: *Computers in Human Behavior* 28.2, pp. 434–443. ISSN: 0747-5632. DOI: http://dx.doi.org/10.1016/j.chb.2011.10.014. URL: http://www.sciencedirect.com/science/article/pii/S0747563211002317.

Le Bon, Gustave (1896). *The Crowd: A Study of the popular mind*. URL: https://archive.org/details/crowdastudypopu00bongoog.

Levmore, S. and M.C. Nussbaum (2010). *The offensive Internet*. Harvard University Press. ISBN: 9780674050891. URL: http://books.google.fr/books?id=Qb543eynQeMC.

Ling, Kimberly et al. (2005). "Using Social Psychology to Motivate Contributions to Online Communities". In: *Journal of Computer-Mediated Communication* 10.4, pp. 00–00. ISSN: 1083-6101. DOI: 10.1111/j.1083-6101.2005.tb00273.x. URL: http://dx.doi.org/10.1111/j.1083-6101.2005.tb00273.x.

Lpez, Claudia and Rosta Farzan (2015). "Lend Me Sugar, I Am Your Neighbor!: A Content Analysis of Online Forums for Local Communities". In: *Proceedings of the 7th International Conference on Communities and Technologies*. C&T '15. Limerick, Ireland: ACM, pp. 59–67. ISBN: 978-1-4503-3460-0. DOI: 10.1145/2768545.2768558. URL: http://doi.acm.org/10.1145/2768545.2768558.

Ma, Xiao, Jeff Hancock, and Mor Naaman (2016). "Anonymity, Intimacy and Self-Disclosure in Social Media". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, pp. 3857–3869. DOI: 10.1145/2858036.2858414. URL: http://doi.acm.org/10.1145/2858036.2858414.

Matias, J. Nathan (2016). "Going Dark: Social Factors in Collective Action Against Platform Operators in the Reddit Blackout". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16.

Santa Clara, California, USA: ACM, pp. 1138–1151. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858391. URL: http://doi.acm.org/10.1145/2858036.2858391.

May, Timothy C. (1992). *The Crypto Anarchist Manifesto*. URL: http://activism.net/cypherpunk/crypto-anarchy.html.

McLeod, Poppy Lauretta (1997). "A Comprehensive Model of Anonymity in Computer-supported Group Decision Making". In: *Proceedings of the Eighteenth International Conference on Information Systems*. ICIS '97. Atlanta, Georgia, USA: Association for Information Systems, pp. 223–234. URL: http://dl.acm.org/citation.cfm?id=353071.353124.

Muller, Michael (2012). "Lurking As Personal Trait or Situational Disposition: Lurking and Contributing in Enterprise Social Media". In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. Seattle, Washington, USA: ACM, pp. 253–256. ISBN: 978-1-4503-1086-4. DOI: 10.1145/2145204.2145245. URL: http://doi.acm.org/10.1145/2145204.2145245.

Nagel, Emily van der and Jordan Frith (2015). "Anonymity, pseudonymity, and the agency of online identity: Examining the social practices of r/Gonewild". In: *First Monday* 20.3. ISSN: 13960466. URL: http://firstmonday.org/ojs/index.php/fm/article/view/5615.

Nelimarkka, Matti, Kai Kuikkaniemi, and Giulio Jacucci (2014). "A Field Trial of an Anonymous Backchannel Among Primary School Pupils". In: *Proceedings of the 18th International Conference on Supporting Group Work*. GROUP '14. Sanibel Island, Florida, USA: ACM, pp. 238–242. ISBN: 978-1-4503-3043-5. DOI: 10.1145/2660398.2660399. URL: http://doi.acm.org/10.1145/2660398.2660399.

Nonnecke, Blair and Jenny Preece (2001). *Why Lurkers Lurk*.

Nonneeke, Blair and Jenny Preece (2000). "Lurker Demographics: Counting the Silent". In: *CHI '00 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2000*. Print, pp. 73–80.

Omernick, Eli and Sara Owsley Sood (2013). "The Impact of Anonymity in Online Communities". In: *International Conference on Social Computing, SocialCom 2013, SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013, Washington, DC, USA, 8-14 September, 2013*, pp. 526–535. DOI: 10.1109/

SocialCom.2013.80. URL: http://doi.ieeecomputersociety.org/10.1109/SocialCom.2013.80.

Oostendorp, Herre van and Ferdy van Varik (2011). "Stimulation of activity in online communities." In: *ECCE*. Ed. by Anke Dittmar and Peter Forbrig. Universittsdruckerei Rostock, Germany / ACM, pp. 163–170. ISBN: 978-1-4503-1029-1. URL: http://dblp.uni-trier.de/db/conf/ecce/ecce2011.html#OostendorpV11.

Ostrom, Elinor, Roy Gardner, and James Walker (1994). *Rules, games, and common-pool resources*. University of Michigan Press.

Park, Deokgun et al. (2016). "Supporting Comment Moderators in Identifying High Quality Online News Comments". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 1114–1125. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858389. URL: http://doi.acm.org/10.1145/2858036.2858389.

Paskuda, Malte and Myriam Lewkowicz (2015). "Anonymous Quorans Are Still Quorans, Just Anonymous". In: *Proceedings of the 7th International Conference on Communities and Technologies*. C&T '15. Limerick, Ireland: ACM, pp. 9–18. ISBN: 978-1-4503-3460-0. DOI: 10.1145/2768545.2768551. URL: http://doi.acm.org/10.1145/2768545.2768551.

— (2016). "Anonymity interacting with participation on a Q&A site". In: *AI & SOCIETY*, pp. 1–13. ISSN: 1435-5655. DOI: 10.1007/s00146-016-0660-9. URL: http://dx.doi.org/10.1007/s00146-016-0660-9.

Phelan, Chanda, Cliff Lampe, and Paul Resnick (2016). "It's Creepy, But It Doesn'T Bother Me". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 5240–5251. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858381. URL: http://doi.acm.org/10.1145/2858036.2858381.

Postmes, Tom, Russell Spears, and Martin Lea (2002). "Intergroup differentiation in computer-mediated communication: Effects of depersonalization." In: *Group Dynamics: Theory, Research, and Practice* 6.1, p. 3.

Preece, Jennifer and Ben Shneiderman (2009). "The reader-to-leader framework: Motivating technology-mediated social participation". In: *AIS Transactions on Human-Computer Interaction* 1.1, pp. 13–32.

Preece, Jenny, Blair Nonnecke, and Dorine Andrews (2004). "The top five reasons for lurking: improving community experiences for everyone". In: *Computers in Human Behavior* 20.2, pp. 201–223.

Rains, Stephen A. (2007). "The Impact of Anonymity on Perceptions of Source Credibility and Influence in Computer-Mediated Group Communication: A Test of Two Competing Hypotheses". In: *Communication Research* 34.1, pp. 100–125. DOI: 10.1177/0093650206296084. eprint: http://crx.sagepub.com/content/34/1/100.full.pdf+html. URL: http://crx.sagepub.com/content/34/1/100.abstract.

Reicher, S D, R Spears, and T Postmes (1995). "A social identity model of deindividuation phenomena". In: *Eur. Rev. Soc. Psychol* 6.

Ren, Yuqing, Robert Kraut, and Sara Kiesler (2007). "Applying common identity and bond theory to design of online communities". In: *Organization studies* 28.3, pp. 377–408.

Ren, Yuqing, Robert Kraut, Sara Kiesler, and Paul Resnick (2012). "Encouraging commitment in online communities". In: *Building successful online communities: Evidence-based social design*, pp. 77–125.

Rocco, E and M Warglien (1996). *Computer mediated communication and the emergence of electronic opportunism. University of Trento working paper.*

Salem, Deborah A., G. Anne Bogat, and Christina Reid (1997). "Mutual help goes on-line". In: *Journal of Community Psychology* 25.2, pp. 189–207. ISSN: 1520-6629. DOI: 10.1002/(SICI)1520-6629(199703)25:2<189::AID-JCOP7>3.0.CO;2-T. URL: http://dx.doi.org/10.1002/(SICI)1520-6629(199703)25:2%3C189::AID-JCOP7%3E3.0.CO;2-T.

Santana, Arthur (2012). "Civility, anonymity and the breakdown of a new public sphere". In:

Sassenberg, Kai and Tom Postmes (2002). "Cognitive and strategic processes in small groups: effects of anonymity of the self and anonymity of the group on social influence." In: *Br J Soc Psychol* 41.Pt 3, pp. 463–80. ISSN: 0144-6665. URL: http://www.biomedsearch.com/nih/Cognitive-strategic-processes-in-small/12419013.html.

Schneider, Hanna et al. (2016). "Understanding the Mechanics of Persuasive System Design: A Mixed-Method Theory-driven Analysis of Freeletics". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 309–320. ISBN:

978-1-4503-3362-7. DOI: 10.1145/2858036.2858290. URL: http://doi.acm.org/10.1145/2858036.2858290.

Sharma, Nikhil et al. (2011). "Emphasizing social features in information portals: Effects on new member engagement". In: *Journal of the American Society for Information Science and Technology* 62.11, pp. 2106–2120. ISSN: 1532-2890. DOI: 10.1002/asi.21624. URL: http://dx.doi.org/10.1002/asi.21624.

Shiue, Yih-Chearng, Chao-Min Chiu, and Chen-Chi Chang (2010). "Exploring and Mitigating Social Loafing in Online Communities". In: *Comput. Hum. Behav.* 26.4, pp. 768–777. ISSN: 0747-5632. DOI: 10.1016/j.chb.2010.01.014. URL: http://dx.doi.org/10.1016/j.chb.2010.01.014.

Shmatikov, Vitaly and Ming-Hsiu Wang (2006). "Measuring relationship anonymity in mix networks". In: *Proceedings of the 5th ACM workshop on Privacy in electronic society*. ACM, pp. 59–62.

Suler, John (2004). "The online disinhibition effect". In: *Cyberpsychology & behavior* 7.3, pp. 321–326.

Tausczik, Yla R., Laura A. Dabbish, and Robert E. Kraut (2014). "Building loyalty to online communities through bond and identity-based attachment to sub-groups". In: *CSCW*, pp. 146–157.

Tsikerdekis, Michail (2013). "The effects of perceived anonymity and anonymity states on conformity and groupthink in online communities: A Wikipedia study". In: *Journal of the American Society for Information Science and Technology* 64.5, pp. 1001–1015. ISSN: 1532-2890. DOI: 10.1002/asi.22795. URL: http://dx.doi.org/10.1002/asi.22795.

Tu, Chih-Hsiung (2002). "The relationship between social presence and online privacy". In: *The Internet and Higher Education* 5.4, pp. 293–318. ISSN: 1096-7516. DOI: http://dx.doi.org/10.1016/S1096-7516(02)00134-3. URL: http://www.sciencedirect.com/science/article/pii/S1096751602001343.

Valacich, Joseph S. et al. (1992). "A conceptual framework of anonymity in Group Support Systems". In: *Group Decision and Negotiation* 1.3, pp. 219–241. ISSN: 1572-9907. DOI: 10.1007/BF00126264. URL: http://dx.doi.org/10.1007/BF00126264.

Wang, Gang et al. (2014). "Whispers in the Dark: Analysis of an Anonymous Social Network". In: *Proceedings of the 2014 Conference on Internet Mea-*

surement Conference. IMC '14. Vancouver, BC, Canada: ACM, pp. 137–150. ISBN: 978-1-4503-3213-2. DOI: 10.1145/2663716.2663728. URL: http://doi.acm.org/10.1145/2663716.2663728.

Wild, Fridolin and Christina Stahl (2007). "Using Latent Semantic Analysis to Assess Social Competence". In: *Mini-Proceedings of the 1st European Workshop on Latent Semantic Analysis in Technology-Enhanced Learning*. Ed. by Wild et al. Open University Netherlands, pp. 3–4.

Wise, Kevin, Brian Hamman, and Kjerstin Thorson (2006). "Moderation, Response Rate, and Message Interactivity: Features of Online Communities and Their Effects on Intent to Participate". In: *Journal of Computer-Mediated Communication* 12.1, pp. 24–41. ISSN: 1083-6101. DOI: 10.1111/j.1083-6101.2006.00313.x. URL: http://dx.doi.org/10.1111/j.1083-6101.2006.00313.x.

Zhu, Haiyi, Sauvik Das, et al. (2016). "A Market in Your Social Network: The Effects of Extrinsic Rewards on Friendsourcing and Relationships". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. Santa Clara, California, USA: ACM, pp. 598–609. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858396. URL: http://doi.acm.org/10.1145/2858036.2858396.

Zhu, Haiyi, Robert Kraut, and Aniket Kittur (2012). "Effectiveness of Shared Leadership in Online Communities". In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. Seattle, Washington, USA: ACM, pp. 407–416. ISBN: 978-1-4503-1086-4. DOI: 10.1145/2145204.2145269. URL: http://doi.acm.org/10.1145/2145204.2145269.

Zhu, Haiyi, Amy Zhang, et al. (2013). "Effects of Peer Feedback on Contribution: A Field Experiment in Wikipedia". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, pp. 2253–2262. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2481311. URL: http://doi.acm.org/10.1145/2470654.2481311.

Chapter 6

Appendix

Résumé : L'influence de l'anonymat sur la participation dans les communautés en ligne

Malte Paskuda

30 août 2016

RESUMÉ

Ce document est une présentation de ma thèse de doctorat sur l'influence de l'anonymat sur la participation dans les communautés en ligne. Le point de départ de ce travail est une observation au cours de la conception d'une plate-forme en ligne pour le soutien social entre aidants informels. J'avais noté que nous ne savions pas quel effet aurait l'option pour les aidants d'être anonyme sur la plate-forme. Ma thèse comporte une revue de la littérature qui est synthétisée dans un modèle qui décrit quel sont les facteurs qui influence la participation en ligne. Nous avons également conduit trois études : Une sur Youtube, dont le système d'identification a changé pour interdire l'anonymat, une sur Quora, où des utilisateurs peuvent choisir de répondre de manière anonyme, et une sur Hacker News, où les utilisateurs d´evoilent plus ou moins leur identité. Ces études nous permettent de montrer que contrairement à ce que dit la littérature, 1) l'anonymat ne conduit pas nécessairement à des discussions impolies, 2) qu'il y a d'autre facteurs que l'anonymat qui ont une influence plus grande sur la participation, et que 3) l'anonymat peut rendre visible l'effet de signaux sociaux, par exemple les attributs de texte comme sa longueur, qui influencent l'appréciation sociale. De plus, nous avons observé que la participation est liée au niveau de détail du profil des utilisateurs, et qu'une forte présence sur le Web par ailleurs peut limiter la participation. Ces résultats permettent de confirmer le modèle "Social Identity of Deindividuation Effects", et le fait que l'anonymat peut avoir une influence positive sur l'esprit de groupe.

INTRODUCTION

Depuis plusieurs décennies, les chercheurs dans le domaine de l'interaction homme-machine (HCI) et du Travail Coopératif Assisté par Ordinateur (CSCW) étudient comment les gens interagissent en utilisant la technologie. Parfois, ces études traitent de questions pour lesquelles d'autres domaines de recherche ont déjà établi des résultats. C'est le cas de l'anonymat.

L'anonymat est souvent perçu comme une qualité négative des environnements en ligne, qui peut conduire à des comportements inappropriés. Certains politiciens ont essayé de supprimer la communications anonyme en ligne, en proposant un enregistrement de chaque utilisateur[1]. En fin, des théories en psychologie comme la désindividuation se sont focalisées sur la façon dont l'anonymat pouvait transformer des groupes de personnes "normales" en une population violente.

Malgré ces travaux, alors que nous participions à la conception d'une plate-forme pour faciliter le soutien social entre aidants informels, nous nous sommes rendus compte que la question de l'anonymat n'était pas facile à résoudre. En effet, malgré les hypothèses négatives, permettre une communications anonyme pourrait avoir des bénéfices importants pour des aidants familiaux : On peut imaginer que utilisateurs seraient plus à l'aise pour discuter de questions sensibles et intimes, comme des troubles du comportement de leurs proches souffrant de la maladie d'Alzheimer par exemple (Salem et al., 1997). Et de manière plus générale, l'option d'être anonyme ne pourrait-elle pas jouer un rôle dans la création d'une identité en ligne (Nagel et Frith, 2015)?

[1] Par exemple https://www.taz.de/!5274217/. Le politician Fischer veux un interdiction de l'anonymat en ligne, car il pense que le qualité des discussions devenir faible. Il dit que c'est un problème si gens crois de pas être responsable pour leur remarques : Anders sah das der damalige Vorsitzende der Enquete-Kommission "Internet und digitale Gesellschaft" Axel Fischer (CDU), der sich für ein "Vermummungsverbot im Internet" aussprach. Fischer argumentierte, dass unter der Möglichkeit sich pseudonymisiert im Netz zu äußern "die Qualität von Diskussionen in Foren und Blogs" leide. Die Anonymität verleite Nutzer zu Äußerungen, die sie hinterher bereuen könnten. Er halte es für bedenklich, dass sich Nutzer durch ein selbst gewähltes Pseudonym vermeintlich jeglicher Verantwortung für Äußerungen entzogen.

REVUE DE LA LITTÉRATURE

La revue de la littérature a abordé plusieurs disciplines. Je me suis focalisé sur des études en HCI, CSCW et en psychologie sociale, tout en prenant connaissance de travaux en sciences politiques et en journalisme.

Il y a deux théories centrales habituellement mobilisées pour expliquer l'effet de l'anonymat sur l'identité et le comportement humain.

La première est la théorie classique de Désindividuation (Postmes et al., 2002), influencée par les théories sur le fonctionnement des groupes du 19è siècle (Le Bon, 1896; Reicher et al., 1995). L'idée centrale est que les normes sociales des individus disparaissent quand ils sont dans un groupe, et ils régressent vers un comportement primitif.

La deuxieme théorie est celle du model "Social Identity Model of Deindividuation Effects" (SIDE) (Reicher et al., 1995). On peut interpréter SIDE comme un théorie de Désindividuation, mais le mécanisme et la conséquence de l'anonymat sont expliquées de manière différente (Cress, 2005). Dans le modèle SIDE, les membres d'un groupe ne perdent pas leur normes sociales, mais ils s'adaptent aux normes en vigueur au sein du groupe. Dans ce cadre, l'anonymat est perçu comme minimisant les différences entre les membres et permet à chaque individu une identification avec les autres plus aisée. Le comportement d'un utilisateur peut donc être influencé par son anonymat, et ce comportement peut être en conflit avec les normes sociales, mais l'effet de l'anonymat n'est pas forcément négatif.

Une étude que nous a semblé importante a été menée par Kilner et Hoadley (2005). Ils ont réussi à observer les évolutions d'un forum entre soldats d'un modèle anonyme vers un modèle exigeant l'identité civile. Ils ont mesuré l'effet de cette évolution sur la participation des soldats, et sur les sujets traités. Les auteurs ont constaté que la participation était plus importante et que la qualité du discours avait augmenté une fois les soldats identifiables.

Une étude comparable a été menée quand le site TechCrunch a changé son système de commentaires, de commentaires anonyme à des commentaires exigeant un identité civil (Omernick et Sood, 2013). Omernick et Sood ont collecté les données de un an avant et un an après le changement. Ils ont analysé ces données en regardant la pertinence des commentaires, leur négativité, et la présence de jurons ou d'insultes. Ils ont combiné ces analyses avec des mesures de participation. Ils ont constaté une augmentation de la qualité des commentaires, et une participation qui avait augmenté sur certains sujets baissé dans d'autres.

En ce qui concerne la participation en ligne, il y a plusieurs approches. Les théories centrales sont la Common Identity Theory et la Common Bond Theory utilisées par Ren et al. qui

fournissent un modèle puissant de ce qui impacte la participation dans une communauté spécifique (Ren et al., 2007). Ren et al. (ibid.) ont identifié par une revue de littérature trois facteurs qui influencent l'identification d'un utilisateur avec un groupe. Ils on noté que la catégorisation sociale, l'interdépendance dane le groupe et la comparaisons entre groupes fortifient l'attachement au groupe, et que l'interaction sociale, le partage d'informations personnelles et les similarités interpersonnelles fortifient les liens interpersonnels au sein du groupe (figure 1).

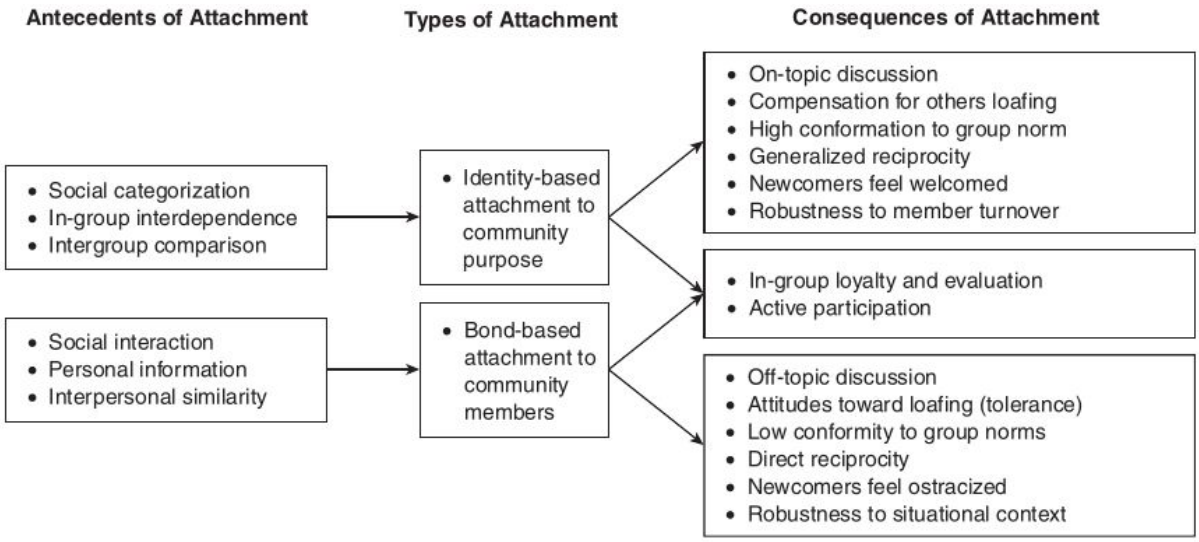

Figure 1 – Facteurs d'attachement de Common Identity Theory et de Common Bond Theory avec leur conséquences, de Ren et al. (2007).

APPROCHE

La première étape de notre approche consistait à faire une revue de littérature détaillée sur l'anonymat et la participation en ligne, puis à se focaliser sur les facteurs pouvant influencer la participation, et parmi eux ceux qui étaient influencés par l'anonymat.

Afin de pouvoir mobiliser ces connaissances, il était nécessaire de la synthétiser dans un modèle qui montre tous les facteurs qui influencent la participation, et qui montre ceux qui sont influencés par l'anonymat. Toutes les connexions dans ce modèle peuvent être envisagées comme des hypothèses.

J'ai utilisé le modèle de synthèse de la littérature (figure 2) pour définir des études me permettant de lier plus finement anonymat et participation en ligne.

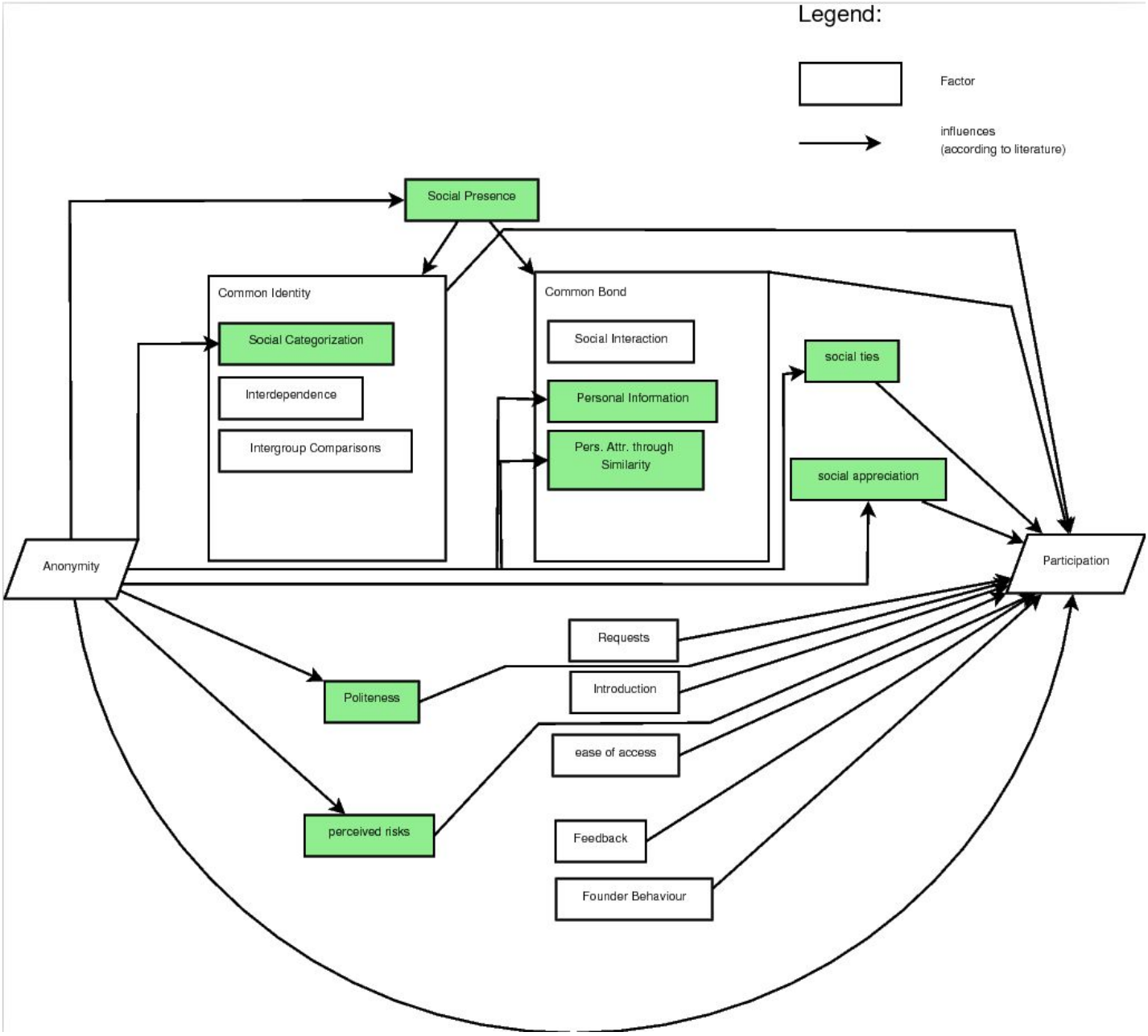

Figure 2 – Le modèle des facteurs qui augmentent la participation, et des facteurs influencés par l'anonymat.

ETUDES

Ce modèle m'a permis de créer trois études sur des communautés en ligne existantes. Le but était à chaque fois d'évaluer le lien prédit par le modèle, mais chaque étude a mis l'accent sur

un aspect différent. La deuxième étude était en effet une réaction à des questions issues de la première, et la troisième une réaction aux questions de la deuxième.

Un concept commun aux trois études est celui de marqueurs. En fait, si on étudie la communication textuelle, on a besoin d'une manière de détecter la présence de facteurs. C'est assez facile pour quelques uns, mais semble impossible pour d'autres. Par exemple l'appréciation sociale : il est possible de la mesurer si une plate-forme dispose d'un système d'évaluation des contributions (*upvotes*), ou d'outils permettant de remercier. Mais mesurer les liens sociaux d'un utilisateur avec d'autres utilisateurs en n'ayant accès à quelques commentaires semble impossible, et on peut dire de même de la perception des risques sociaux. De ce fait, pour chaque facteur que nous avons voulu mesurer, il a été nécessaire de trouver des marqueurs visibles dans le texte. Ces marqueurs dépendent bien sûr de la plate-forme étudiée. Cela signifie que nos marqueurs changent dans chacune de nos trois études, et que chaque étude repose sur un modèle simplifié avec les facteurs appropriés.

Etude YouTube

Nous avons choisi d'étudier la plate-forme YouTube parce qu'en novembre 2013, Google a fusionné le système de commentaires de YouTube et Google+. Avant cette date, les utilisateurs étaient libres de choisir un identifiant, mais après cette date, ils ont etaient forcé d'utiliser leur identité civile (ou tout du moins l'identité qu'ils avaient utilisée pour créer leur compte google). Il nous a donc semblé intéressant de comparer les commentaires sur YouTube avant ce changement et après ce changement.

Le modèle simplifié que nous avons donc utilisé pour cette étude sur YouTube inclut la politesse, la comparaison entre groupes et l'interaction sociale (figure 3).

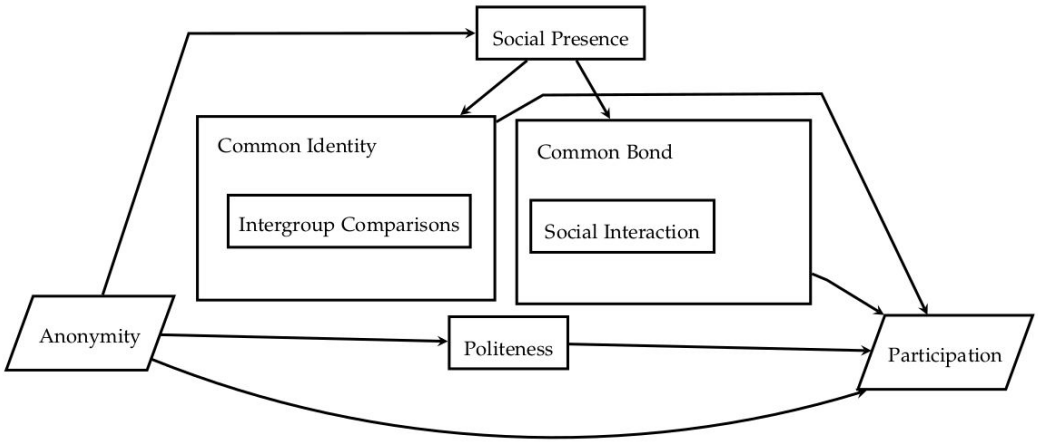

Figure 3 – Modèle de facteurs de participation pour Youtube.

Après le changement, nous avons constaté qu'il y avait plus de commentaires, que ceux ci soient polis ou impolis (tableau 1), un χ^2-test donne $p < 0.01$.

	Polite	Neutral	Rude
Before	133 (3%)	2838 (92%)	155 (5%)
After	32 (5%)	534 (84%)	81 (11%)

Tableau 1 – Evolution de la politesse des commentaires

La comparaison inter-groupes était un peu plus élevée après le changement (tableau 2), mais cette différence n'est pas significative (t-test).

Group	mean	sd	median	n
Before	0.1628	0.5885	0	3126
After	0.2365	1.3417	0	647

Tableau 2 – Evolution du nombre de comparaisons

Le nombre de réponses après le changement est élevé (tableau 3) et signifiant par t-test avec $p < 0.01$.

Group	mean	sd	median	n
Before	0.0067	0.1171	9	3126
After	0.4791	2.3598	0	647

Tableau 3 – Evolution du nombre de réponses

Cette étude nous a permis de générer deux hypothèses : (1) Si les commentateurs sont anonymes, ils postent moins de commentaires polis et moins de commentaires impolis; (2) Si les commentateurs sont anonymes, il y a moins d'interaction.

Quora

Après avoir constaté dans l'étude sur YouTube que l'anonymat n'avait pas eu l'effet escompté et que son effet était plus faible que ce que l'on aurait pu imaginer, l'étape suivante consistait à

chercher un environnement en ligne dans lequel d'autres facteurs que l'anonymat auraient pu influencer la participation. La plate-forme de questions/réponses Quora nous a paru répondre à ce critère, car ses utilisateurs peuvent choisir de poser des questions et de répondre à des questions, avec leur profil ou en restant anonyme. Cela nous permet donc de regarder la différence entre les posts anonymes qui ont pu être postés au même moment, dans le même environnement, contrairement à l'étude précédente dans laquelle un biais pouvait être le changement de culture des utilisateurs après le changement dans le système de commentaires.

Pour cette étude sur Quora, le modèle simplifié comporte les facteurs suivants : politesse, comparaisons entre groupes et appréciation sociale (figure 4).

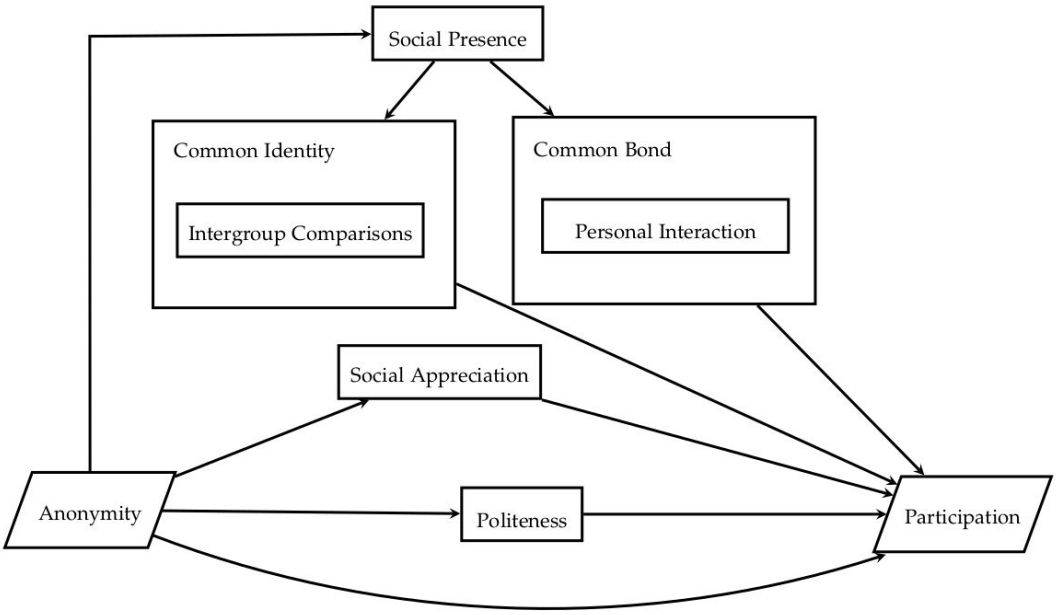

Figure 4 – Modèle de facteurs pour l'étude sur Quora.

Nous avons constaté une légère différence de comparaisons entre les groupes entre les messages dont les utilisateurs étaient identifiés et les anonymes (figure 5).

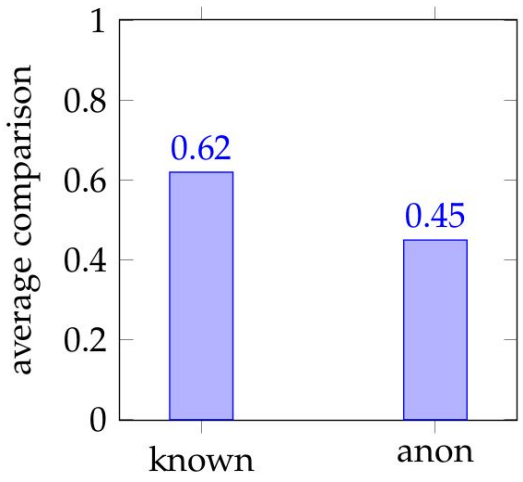

Figure 5 - Comparaison du nombre moyen de comparaison entre groupes

Un t-test montre que la différence entre le nombre de messages reçus n'est pas signifiant (figure 6).

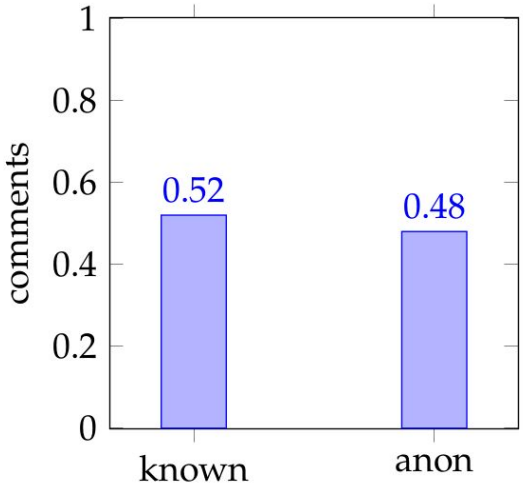

Figure 6 – Moyenne de réponses reçues.

Le nombre moyen d'upvotes (votes positifs) reçu n'était pas significativement différent (figure 7).

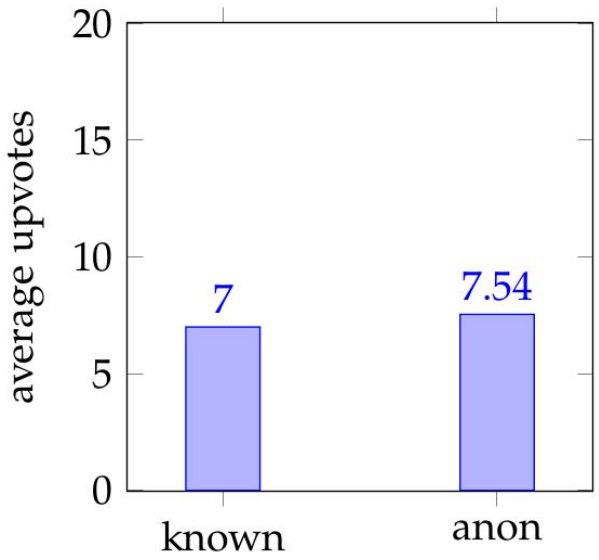

Figure 7 – Nombre moyen de upvotes reçus

En revanche, nous avons trouvé une corrélation de Pearson de r = 0.384 avec une marge de liberté de 286 ($p < 0.01$) entre la longueur d'une réponse et ses upvotes si le réponse est postée de manière anonyme (figure 8).

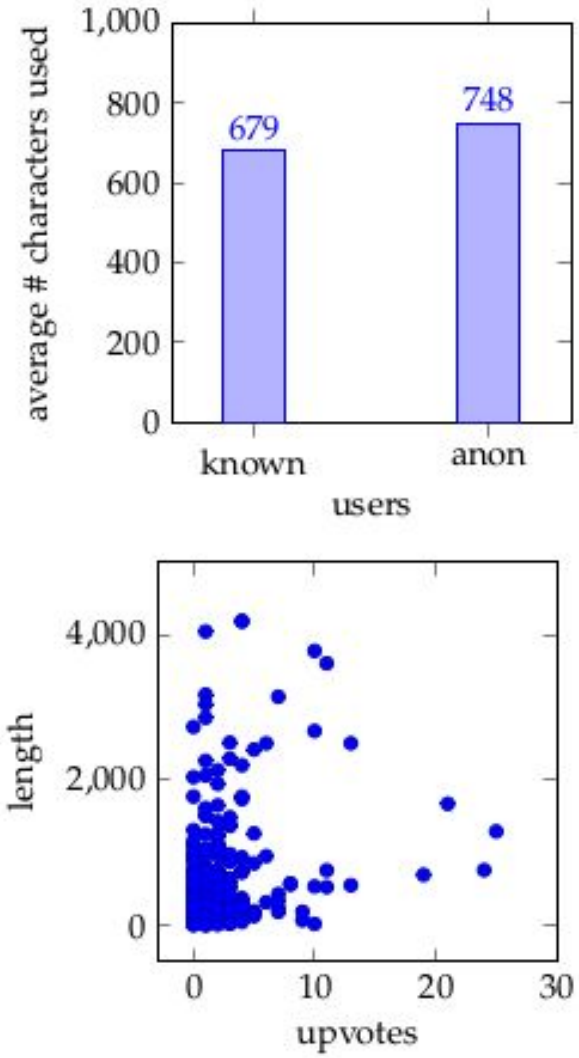

Figure 8 –Corrélation entre longueur de réponses anonyme et upvotes

Quora nous a permis d'étudier un environnement dans lequel des textes sont postés à la fois par des utilisateurs anonymes et des utilisateurs identifiés. Contrairement à l'étude sur YouTube que nous avions conduite précédemment, nous avons constaté peu de différences entre ces textes. Sur YouTube, l'évolution principale était le niveau de politesse et l'augmentation d'interaction sociale. Sur Quora, la seule différence constatée est la corrélation entre la taille des réponses et les votes pour les réponses postées anonymement.

Hacker News

L'objectif de l'étude sur Hacker News était de regarder plus précisément le rôle de facteurs d'identité et de signaux sociaux. En partant des upvotes des commentaires, j'ai essayé de regarder si l'appréciation sociale d'un commentaire est, sur cette plate-forme influencée de manière significative par les de facteurs d'identité de son auteur. Ainsi, cette étude ne s'est pas focalisée directement sur le modèle, mais seulement sur la relation entre l'anonymat (défini par des degrés d'identité) et l'appréciation sociale.

Sur Hacker News, nous avons proposé un score d'identité basé sur des facteurs d'identité et mesuré par Bayes :

$$f(x) = \begin{cases} x * -0.5 & \text{si pseudonyme} \\ \dfrac{x+100}{2} + 50 & \text{so identité civile} \end{cases}$$

Nous avons appliqué ce calcul à 12000 utilisateurs (figure 9).

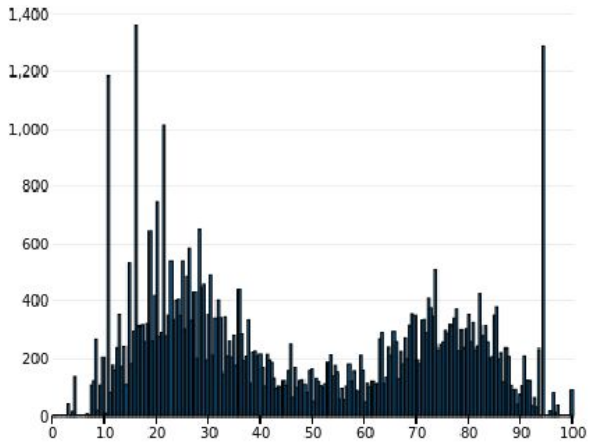

Figure 9 – Histogramme des scores d'identité

Cette étude nous a montré que les upvotes n'étaient pas fortement influencés par les facteurs d'identité (tableau 4).

R	R²	R² Adj	Std. Error R
0.140	0.020	0.019	6.459

Tableau 4 – test Anova des upvotes

Pour les suggestions, au contraire, la corrélation avec le score d'identité était plus importante. (tableau 5).

R	R^2	R^2 Adj	Std. Error R
0.328	0.107	0.107	267.179

Tableau 5 – test Anova des suggestions

La corrélation avec le nombre de commentaires était encore plus forte (tableau 6).

R	R^2	R^2 Adj	Std. Error R
0.538	0.290	0.290	2420.671

Tableau 6 – test Anova des commentaires

Cette étude nous a également montré qu'un utilisateur qui poste plus des suggestions est probablement un auteur dont le profil est davantage décrit (figure 10).

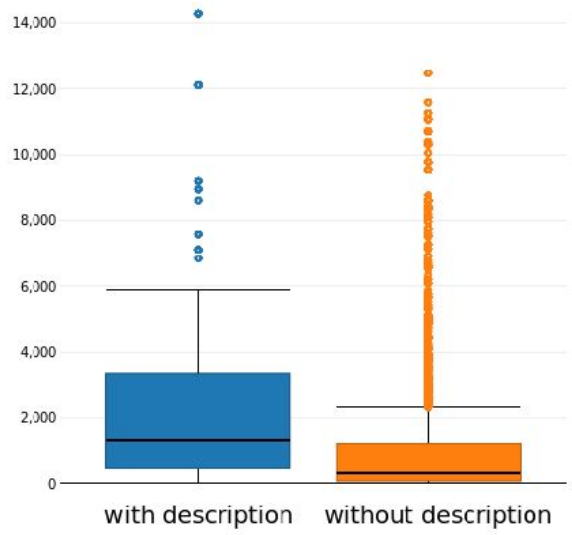

Figure 10 Nombre de commentaires pour les utilisateurs avec et sans description dans leur profil.

Notre but dans cette étude était de mesurer l'effet d'un nombre de facteurs d'identité sur l'appréciation sociale et la participation, dans la continuité des travaux existant (dont notre étude sur YouTube) qui présageaient un effet. Notre analyse de 50.000 commentaires de Hacker News, nous a permis de constater une corrélation significative entre les facteurs d'identité et le nombre d'upvotes reçus par commentaire, et une corrélation entre les facteurs d'identité et le nombre de posts et de commentaires.

CONCLUSION

Ce travail de thèse nous a permis de faire émerger de nouvelles connaissances sur l'effet de l'anonymat sur la participation en ligne, grâce à la création d'un modèle et trois études sur des plate-formes existantes.

La littérature existante m'avait donné l'impression que l'anonymat avait un effet négatif, et conduisait à une impolitesse, une baisse de la qualité de discussions et moins d'appréciation positive pour les commentaires anonymes (Kilner et Hoadley, 2005; Omernick et Sood, 2013; Rains, 2007; Santana, 2012). Mais ce n'est pas ce que nos études ont montré.

Nous avons en effet constaté sur Hacker News une corrélation faible entre les informations fournies sur l'identité et la participation. Et l'anonymat n'engendre pas sur cette plate-forme de baisse de l'appréciation sociale, ou de baisse de la qualité des échanges.

Nous avons constaté des effets de l'anonymat. Sur Quora, l'appréciation sociale des réponses anonymes est corrélée avec leur longueur. Sur Hacker News, à la fois les réponses anonymes et non anonyme ont eu cet corrélation. Si l'appréciation est en effet corrélée avec la longueur d'un réponse, d'autres facteurs que l'identité ont dû influencer l'appréciation quand l'utilisateur n'était pas anonyme. La longueur du texte peut être perçue un signal social (Wise et al., 2006), avec les réponses les plus longues considérées comme de meilleure qualité. On peut interpréter la suppression des signaux sociaux comme un point positif et argumenter comme Cress (2005) que supprimer les informations personnelles comme les images amène les utilisateurs à s'identifier davantage aux membres anonymes. La longueur des commentaires serait donc ainsi un facteur objectif de la qualité des commentaires, un facteur qui était auparavant masqué par des éléments d'identité comme un nom d'utilisateur ou une image de profil perçus négativement.

Le travail de cette thèse pourrait donner lieu à deux types d'études complémentaires : (1) Réaliser des expérimentations en situations contrôlées pour vérifier certaines hypothèses, (2) Etudier des communautés en ligne permettant d'isoler l'effet de l'anonymat, en comparant deux groupes, un avec des contributeurs anonyme et un autre avec des contributeurs utilisant leur

identité civile. C'était d'ailleurs notre idée pour le projet AAL TOPIC, mais que nous n'avons pas pu mener à bien faute d'un nombre d'utilisateurs suffisants dans la période de notre recherche.

REFERENCES

Cress, Ulrike (2005). "Why Member Portraits Can Undermine Participation". In : Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years! CSCL '05. Taipei, Taiwan : International Society of the Learning Sciences, p. 86–90. isbn : 0-8058-5782-6. url : http://dl.acm. org/citation.cfm?id=1149293.1149305.

Kilner, Peter G. et Christopher M. Hoadley (2005). "Anonymity Options and Professional Participation in an Online Community of Practice". In : Proceedings of th 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years! CSCL '05. Taipei, Taiwan : International Society of the Learning Sciences, p. 272–280. isbn : 0-8058-5782-6. url : http://dl.acm. org/hcitation.cfm?id=1149293.1149328.

Le Bon, Gustave (1896). The Crowd: A Study of the popular mind. url : https://archive.org/details/crowdastudypopu00bongoog.

Nagel, Emily van der et Jordan Frith (2015). "Anonymity, pseudonymity, and the agency of online identity: Examining the social practices of r/Gonewild". In : First Monday 20.3. issn : 13960466. url : http://firstmonday.org/ojs/index.php/fm/article/ view/5615.

Omernick, Eli et Sara Owsley Sood (2013). "The Impact of Anonymity in Online Communities". In : International Conference on Social Computing, SocialCom 2013, SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013, Washington, DC, USA, 8-14 September, 2013, p. 526–535. doi : 10.1109/SocialCom.2013.80. url : http://doi.ieeecomputersociety. org/10.1109/SocialCom.2013.80.

Postmes, Tom, Russell Spears et Martin Lea (2002). "Intergroup differentiation in computer-mediated communication: Effects of depersonalization." In : Group Dynamics: Theory, Research, and Prac- tice 6.1, p. 3.

Rains, Stephen A. (2007). "The Impact of Anonymity on Perceptions of Source Credibility and Influence in Computer-Mediated Group Communication: A Test of Two Competing Hypotheses". In : Communication Research 34.1, p. 100–125. doi : 10.1177/0093650206296084. eprint : http://crx.sagepub.com/content/34/1/100.full. pdf+html. url : http://crx.sagepub.com/content/34/1/100. abstract.

Reicher, SD, R Spears et T Postmes (1995). “A social identity model of deindividuation phenomena”. In : Eur. Rev. Soc. Psychol 6. Ren, Yuqing, Robert Kraut et Sara Kiesler (2007). “Applying common identity and bond theory to design of online communities”. In : Organization studies 28.3, p. 377–408.

Salem, Deborah A., G. Anne Bogat et Christina Reid (1997). “Mutual help goes on-line”. In : Journal of Community Psychology 25.2, p. 189–207. issn : 1520-6629. doi : 10.1002/(SICI)1520-6629(199703) 25:2<189::AID-JCOP7>3.0.CO;2-T. url : http://dx.doi.org/10.1002/(SICI)1520-6629(199703)25:2%3C189::AID- JCOP7%3E3.0.CO;2-T.

Santana, Arthur (2012). “Civility, anonymity and the breakdown of a new public sphere”.

Wise, Kevin, Brian Hamman et Kjerstin Thorson (2006). “Moderation, Response Rate, and Message Interactivity: Features of Online Communities and Their Effects on Intent to Participate”. In : Journal of Computer-Mediated Communication 12.1, p. 24–41. issn : 1083-6101. doi : 10.1111/j.1083-6101.2006.00313.x. url : http://dx.doi.org/10.1111/j.1083-6101.2006.00313.x.

Malte PASKUDA

Doctorat : Ingénierie Sociotechnique des Connaissances, des Réseaux et du Développement Durable

Année 2016

L'influence de l'anonymat sur la participation dans les communautés en ligne

Cette thèse de doctorat porte sur l'influence de l'anonymat sur la participation dans les communautés en ligne. Le point de départ de ce travail est une observation au cours de la conception d'une plate-forme en ligne pour le soutien social entre aidants informels. J'avais noté que nous ne savions pas décider si les aidants devaient pouvoir être anonymes sur la plate-forme ou non, et quel en serait l'effet. Ma thèse comporte une revue de la littérature qui est synthétisée dans un modèle qui décrit quel sont les facteurs qui de participation en ligne qui pourraient être influencés par l'anonymat. Nous avons conduit trois études : Une sur Youtube, dont le système d'identification a changé pour ne plus permettre de poster des commentaires de façon anonyme, une sur Quora, où les utilisateurs peuvent choisir de répondre aux questions de manière anonyme ou non, et une sur Hacker News, où les utilisateurs peuvent choisir de dévoiler plus ou moins leur identité. Ces études nous permettent de montrer que, contrairement à ce que dit la littérature, 1) l'anonymat ne conduit pas nécessairement à des discussions impolies, 2) qu'il y a d'autre facteurs que l'anonymat qui ont une influence plus importante sur la participation, et que 3) l'anonymat peut révéler d'autres facteurs qui ont un effet sur la participation, comme la longueur du texte, qui a un effet sur l'appréciation sociale. Ces résultats permettent de confirmer le modèle "Social Identity of Deindividuation Effects", et le fait que l'anonymat peut avoir une influence positive sur l'esprit de groupe.

Mots clés : anonymat - réseaux sociaux (Internet) - participation sociale - identité numérique.

The Influence of Anonymity on Participation in Online Communities

This work presents my PhD thesis over the influence of anonymity on participation in online environments. The starting point was the observation made during the design of an online platform for informal caregivers, where I realized that it was unknown to us which practical effects an anonymous identity would have on the participation. This work contains the subsequent literature review, which was synthesized into a model showing which participation factors might be influenced by anonymity. We conducted three studies on existing online environments: One on Youtube, where there was a change in the comment system forbidding anonymous comments, one on Quora, where users can opt to answer questions anonymously, and one on Hacker News, where users choose how many identity factors they present and which name they use. The result of these studies are that, contrary to what the literature would suggest, 1) anonymity did not result to impolite and uncivil discussion, and 2) other factors than anonymity have a stronger influence on participation, and that 3) anonymity can make the effect of social signals visible, e.g. text properties like length which influences social appreciation. Additionally, we observed that participation is linked to profile completeness, and that an established web presence elsewhere limits participation. The implications of these results are a confirmation of the Social Identity Model of Deindividuation Effects, in its interpretation that anonymity can have positive effects on group identity.

Keywords: anonymity - online social networks - social participation - online identities.

Thèse réalisée en partenariat entre :

anrt association nationale recherche technologie

CNSA Caisse nationale de solidarité pour l'autonomie

utt université de technologie Troyes

Ecole Doctorale "Sciences et Technologies"