Knowledge Evolution in Social Media: Understanding the Effects of Website Designs on the Selection of User-Generated Content

By:

Gabriela Morales Martinez

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Social Sciences
Management School

August 2018

Dedicated to my mentor and best friend
It's really nice to be your daughter, dad

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr. Dermot Breslin and Prof. Paul Latreille. Their comments, guidance and critical eye have been central in developing the thesis until this point. Specifically, I would like to thank Dermot for introducing me to novel methodologies, and Paul for being encouraging and believing that my research was worth pursuing. Moreover, this thesis would not have been possible without Dr. Paul Denny, who agreed to make changes on PeerWise, and Dr. Jon Burchell, who allowed me to use his module for data collection. Further, I would like to thank Dr. Mark Taylor, Dr. Basile Marquier, Dr. Umberto Esposito and Prof. Matthew Jackson, for their advice and feedback regarding data analysis; as well as Dr. Elaine Toms, Prof. Andrew Tylecote and Dr. Mario Doria, for reading and commenting on various sections of this research. Lastly, I would like to thank Mexico's *National Council for Science and Technology (CONACTY)*, which granted and sponsored my doctoral studies in the UK.

The past few years would not have been the same without my officemates: Liam, Sam, Rob, Mariana, Lorian, Caro L., Bavs, Ella, Zati, Caro V., Nour, Nik, Tobi, Deepak & Albert. I have heard that doing a PhD can be a lonely experience, but this was far from my reality. Actually, I could have used some 'me time'! Thank you for all the laughs, coffee breaks, delicious treats from your countries, and even for that one camping trip where we realised it was better for PhDs to stay indoors! Special thanks to Mariana, with whom I shared an office for three years, for all her feedback and encouragement. You made working during weekends seem like fun, *te quiero, chiquis!* I would also like to thank my friends Rose, Kelly, Ale and Nejhme. Having such witty, kind-hearted women like you in my life is a source of joy and empowerment.

Most importantly, I would like to thank my family. Thanks to my parents, Martha and Guillermo, for always being loving and encouraging. I am forever in-debt with life for giving me such kind, smart role models. Also, many thanks to my *Martinitos*: Silvia, David, Toty, Raúl and Adrián, for their every-day love and support. Finally, I would like to thank Umberto. The best thing about coming to the UK to study was meeting you. Thanks for keeping me going when it was tough. The end of this thesis marks the start of our life together, and I could not be happier about both events! *Ti amo, 'esposito'!*

ABSTRACT

Purpose – This thesis aims to understand how the designs of social media sites affect the transmission of user-generated content, through the impact that different set-ups have on identity, groups and relationships, and consequently on the choices of users.

Design/Methodology – A quasi-experiment is set up, consisting of three experimental conditions on an existing online platform. The research uses a mixed- and multiple-method approach where the primary sources of data are online interactions and a questionnaire that, *inter alia*, mapped participants' personal networks, with focus groups used to obtain more in-depth information about the choices made by participants. The examination comprises descriptive and inferential statistics, social network, clustering, sequence, and thematic analyses. Notably, sequence analysis is applied for the first time to the study of social media and ratings.

Findings – A key result is that the different designs of websites act as a frame to the content shared, i.e. online users make different choices depending on how the information is presented. Specifically, different designs affect choices because of: 1) the impact that groups and relationships have on identity management; 2) the type and strength of group-biases; and 3) transmission errors. Most importantly, the online presence of individuals' real-life relationships affects how content is perceived and evaluated.

Contributions – This research adds to theory by conceptualising the process of variation, selection and retention of knowledge in social media and creating a model to study selection through choice-making. Also, it contributes by increasing the understanding of how certain website designs can increase/decrease transmission errors and biases, hence affecting the evolution of knowledge and the 'wisdom of the crowd'. Regarding methodology, this thesis contributes by conducting one of the most complete studies ever performed in online environments, combining different methods of data collection and analysis. As regards practice, the research identifies important design considerations for website developers. Further, concerning policy, the study presents a reflection on frames as ethical acts, and outlines a number of questions that should serve as a basis for debate among policymakers.

Future research – The thesis concludes by outlining seven possible lines of research. This work is expected to trigger the interest of scholars, practitioners, and policymakers in understanding the relevance of appropriate designs for social media sites.

LIST OF ABBREVIATIONS

CC	Content community
CMC	Computer-mediated communication
eWOM	Electronic word-of-mouth
LG-n	Numbered literature gap
MM	Multilevel modelling
NVC	Non-verbal communication
OB-n	Numbered objective
RQ-n	Numbered research question
SA	Sequence analysis
SIDE	Social identity model of deindividuation effects
SNA	Social network analysis
SNS	Social networking sites
UGC	User-generated content
VSR	Variation, selection, and retention (mechanisms)

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Importance and focus of the research

At the time of writing this thesis, the number of internet users is 3.92 billion (Internet Live Stats 2018), slightly over half of the world's population (Worldometers 2018). Moreover, as of April 2018, 2.2 billion of these online users were using Facebook, the world's biggest social media site (Statista 2018a), which makes the population of Facebook more than five times greater than the inhabitants of Europe (PopulationPyramid 2018)[1]. People use social media for entertainment purposes, to do business, buy and sell products, and to communicate (Correa et al. 2010; Hughes et al. 2012). Further, social media is known to have a massive impact on how online users behave, search for information, form communities, build and sustain relationships, and how they generate, modify and share content across sites and devices (Kietzmann et al. 2012). However, there seems to be very little understanding regarding whether, and how, social media sites might affect the transmission of information. For instance, the recent news about *Cambridge Analytica* suggested that social media might have been partly responsible for worldwide political and socio-economic decisions, such as *Brexit* in the UK, and the 2016 presidential elections in the US, together with over a hundred other election campaigns in over thirty countries (Ghoshal 2018; Greenfield 2018).

There are several elements that might affect the sharing of information. Yet a component that seems to have received little attention – from scholars, practitioners, and policymakers – is the design of these sites. As will be explained in the following chapter, to date there are only two studies which compare the set-ups of different social media sites (e.g. Brandtzæg et al. 2009; Li 2017). In addition, it has been argued that there should be more research focusing on the designs of different online platforms (Kietzmann et al. 2012). Accordingly, the purpose of this thesis is to determine whether – and how – certain aspects of the designs of social media sites might enable, restrain or affect the transmission of user-generated content (UGC). A particular focus is placed on the impact that these aspects – user profiles and rating scales – have on identity, groups and relationships, and consequently on the choices of users.

[1] In 2015, when this research was starting, the number of Facebook users was only double the population of Europe.

In regards to choices, this research adopts an evolutionary perspective. Evolutionary narratives explain the process of information transmission in terms of the mechanisms of variation, selection and retention (VSR). Information varies in the form of beliefs; selection takes place when a person acquires these beliefs from others; and retention is present as long as there is no more variation of that particular belief (Weick 1969). In the context of social media and UGC, variation can be seen as the pool of information that individuals generate within social networks, where they communicate their beliefs through content such as text, audio, video, symbols or emoticons. The selection part would then be when an individual or groups of individuals within a network choose a particular piece of information. Finally, the retention mechanism takes place when the people who selected the information retain the conveyed belief. This research focuses on the selection mechanism, that is, on the way online users choose content from other members of the network, studied through the ratings of users.

However, in the era of social media where online data is virtually unlimited and there is an unprecedented number of people interacting, it would be impossible for users to gather and compare the content posted by all online members in making a choice. Therefore, in order for people to obtain information faster and make quicker decisions, they rely on heuristic principles (i.e. 'rules of thumb') that simplify but may also bias the information they acquire (Kahneman 2003; Tversky & Kahneman 1981). Likewise, due to the amount of information that is being produced and consumed by users, online sites have adopted designs that enable heuristics, simplifying the transmission of information, but potentially also biasing the data users receive (Park & Nicolau 2015). In this regard, the different set-ups that websites adopt can affect, among other elements: the amount of self-presentation that users can have through their profiles, the type and length of media that can be shared on the site, the way in which people can evaluate the content of others through the use of rating scales, and the reach that users' comments can have among the whole network. The present research studies the effects that user profiles and rating scales have on the choices of users.

Figure 1.1 presents a general conceptual model of this thesis. As can be seen, the focal point of this research is the selection mechanism, which is investigated through the choice-making process and the biases that take place when individuals within groups make choices. Further, selection is studied in the context of three 'building blocks' of social media: identity, groups and relationships (Kietzmann et al. 2012). Lastly, these concepts are analysed within the setting of social media and its different designs regarding different user profiles (anonymous versus identifiable) and rating scales (likert versus dichotomous).

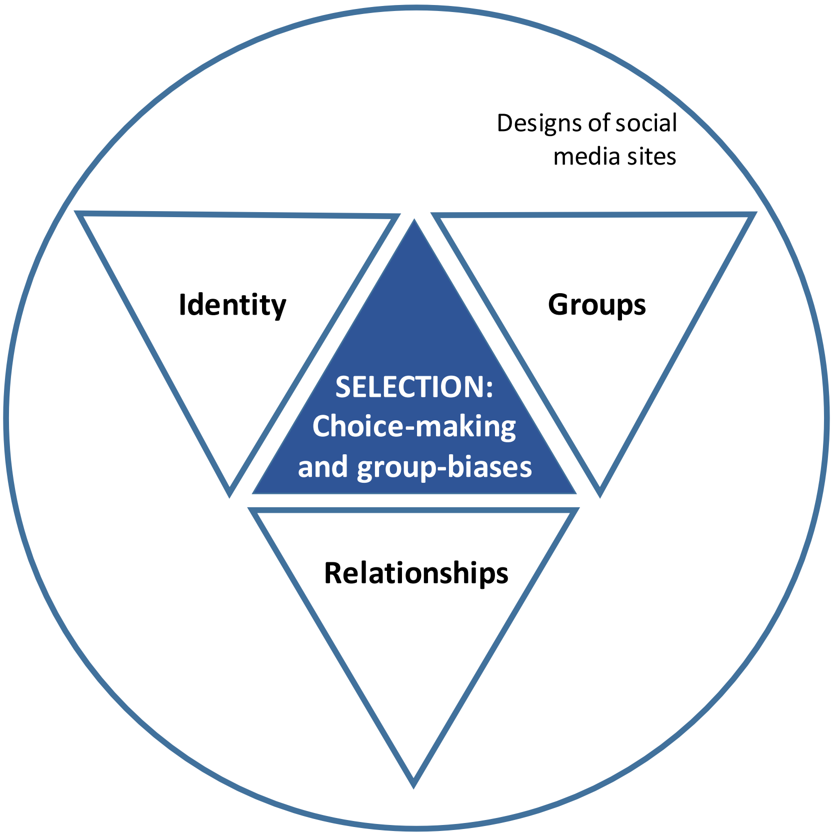


Figure 1.1 - General conceptual model of the thesis

Regarding identity, this thesis builds on the argument that, when a person is in the presence of others, they try to manage the impressions they give (Goffman 1959). Therefore, identity is not an individual attribute, but instead a product of socialisation (Goffman 1959; Altheide 2000). For this thesis, identity management comprises the study of two elements of the classification of social media: self-disclosure and self-presentation (Kaplan & Haenlein 2010). Self-presentation, defined as the way users present themselves in cyberspace, is studied by contrasting two types of user profiles on an educational platform: anonymous and identifiable. Self-disclosure – the revelation of personal information such as likes and dislikes – is investigated through the comparison of two rating scales deployed on the same platform: likert and dichotomous.

Concerning groups, in a similar way to Goffman (1959; 1963), social psychologists have argued that people's identities are subject to groups and thus vary on a continuum between the individual and the collective (Tajfel 1974; Tajfel 1978; Tajfel & Turner 1979). Further, individuals cluster into groups whenever differences arise (e.g. gender, race, political views, socioeconomic status) and tend to favour the *ingroup* as opposed to the *outgroup* (Tajfel 1974). Therefore, it is argued here that in the online space, groups have an impact on the way users self-present and the information they disclose about themselves (i.e. their likes and dislikes). Moreover, the fact that individuals gather in groups and favour similar members, has an impact on the content that they access, or more specifically, from whom they acquire content. Based on this logic, the thesis differentiates between the sharing of content with the entire and with personal networks (i.e. users' outgroups and ingroups).

Further, in terms of relationships, it has long been known that the way in which individuals interact has an impact on their sharing of advice and information (Granovetter 1973). Likewise, research has shown that a person's information environment resides mostly in their social connections (Cross, Parker, et al. 2001). However, not all relations have the same strength and, therefore, users do not present themselves or share the same content with all of their acquaintances. Thus, based on these arguments, the thesis further differentiates within people's personal networks, through the strength of their ties (Granovetter 1973; 1983).

1.2 Aim and Research Questions

This thesis aims to understand how the designs of social media websites might enable, restrain or otherwise affect the transmission of UGC, through the impact that different set-ups have on identity, groups and relationships; and consequently, on the choices of users. The aim is tackled by comparing two designs regarding user profiles (i.e. anonymous and identifiable) and then by contrasting two types of rating scale (i.e. likert and dichotomous). Moreover, the research examines whether these elements affect the choices of users through the study of ratings, emphasising those that are given to the whole network as opposed to those given among individuals' personal networks (i.e. ingroups).

Specifically, the thesis poses an overall and three specific research questions:

- *RQ-Overall:* How is the transmission of UGC affected by the different designs (i.e. frames) adopted by social media websites?
- *RQ-1:* How are the choices of users affected by different levels of self-presentation occurring from diverse user profiles?
- *RQ-2:* How are the choices of users affected by different levels of self-disclosure happening due to the use of distinct rating scales?
- *RQ-3:* How are the choices of users affected by the online presence of their personal networks?

The proposed model, which combines all the concepts presented so far and highlights the posed research questions, is as follows:

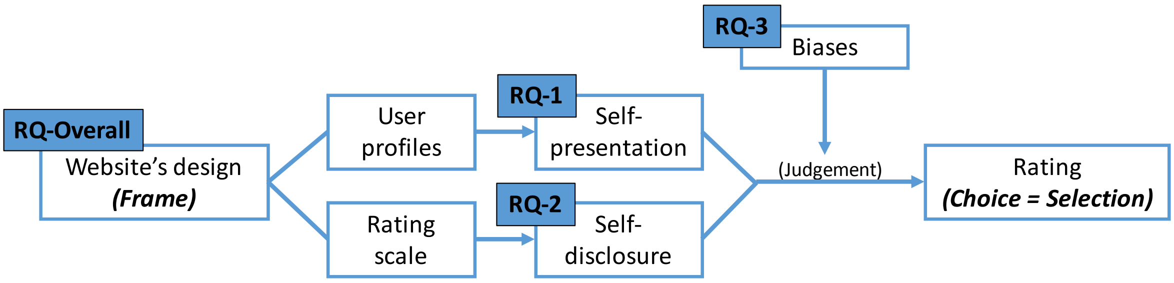


Figure 1.2 - Conceptual model with research questions

In order to address the posed research questions, three quasi-experiment were set-up. These comprised three groups of participants experiencing different online conditions. The chosen website to conduct the quasi-experiments was *PeerWise*, an educational platform where students author and post multiple-choice questions and later answer, rate, and comment on the questions posed by their peers. In order to test different website designs, *PeerWise* was modified with the help of its creator. The data collected comprised approximately 200,000 online interactions, 400 questionnaires, and 6 focus groups. Data were generated over the course of three years by the almost 1,000 participants who took part in the quasi-experiments.

1.3 Thesis structure

The outline of the remainder of the thesis is as follows. Chapter 2 presents a literature review of the key areas and theoretical constructs that underpin the study. It opens with a description of social media and the classification of sites presented by Kaplan & Haenlein (2010). The building blocks of social media are discussed (Kietzmann et al. 2012), with a focus on those of identity (e.g. Goffman 1959; Reicher et al. 1995; Zhao et al. 2008), groups (e.g. Tajfel & Turner 1979), and relationships (e.g. Granovetter 1973; 1983; Cross, Parker, et al. 2001). Subsequently, the transmission of UGC is explained from an evolutionary perspective, and the mechanisms of variation, selection, and retention are outlined (Campbell 1960; Weick 1969; Mesoudi 2011). This is followed by a more detailed focus on the selection of UGC, which forms the core issue of this thesis. Selection is explained through choice-making, drawing on the concepts of bounded rationality (Simon 1955; 1979) and frames (Tversky & Kahneman 1981; Kahneman 2003), and by making an analogy of choices and online ratings (e.g. Riedl et al. 2010; 2013). Thereafter, group-biases are explained in relation to their effect on choices (Richerson & Boyd 2005; Mesoudi 2011). Lastly, the chapter presents the resulting theoretical framework deriving from this synthesis of ideas and from which the model, aim, research questions and objectives the study follow.

Chapter 3 describes the methodology adopted to address the research questions and objectives. This chapter starts by outlining the philosophical underpinnings of the research, which adopts a post-positivist standpoint (e.g. Guba & Lincoln 1994; Miller 2005; Robson 2011; Trochim et al. 2016). Then, the mostly used methodologies in online environments are discussed, highlighting their advantages and shortcomings. Further, the quasi-experimental design is explained (Campbell & Ross 1968), together with the mixed- and multiple-methods approach to data collection (Bryman & Bell 2011). Next, the process of data analysis is described, together with the following methods: descriptive and inferential statistics, social network, clustering, sequence, and thematic analyses.

Research findings are presented in Chapters 4 to 7. Chapter 4 responds to the first research question, testing different levels of self-presentation on ratings by comparing anonymous and identifiable user profiles. Next, Chapter 5 addresses the query concerning different levels of self-disclosure on ratings, by contrasting likert and dichotomous scales. Chapter 6 examines the third research question, investigating whether users conform differently to the whole than to their personal networks. The last chapter of findings, Chapter 7, tackles the overall research question by making a holistic comparison among the three quasi-experimental conditions and deepens the analysis on the study of conformity.

Chapter 8 presents the discussion of the key findings of the research in the context of the reviewed theories, with an emphasis on the '*what*' (i.e. the factors that should be considered in the explanation of the issue of interest), the '*how*' (i.e. the relationships among the factors), and the '*why*' (i.e. the psychological or social dynamics that explain the selection of factors) (Whetten 1989). Specifically, it reveals an enhanced conceptual model for the study and reflects on how website designs impact on the choices of users. Further, it presents a discussion on how knowledge evolution takes place in online environments, and how different website set-ups affect the variation, selection, and retention mechanisms. Finally, Chapter 9 concludes the thesis by summarising its contributions to theory and methods, outlining the implications for practice and policy, and indicating the limitations of the study and suggested paths for future research.

Taken together, this thesis combines knowledge from different disciplines and merges them to explain how and why people's beliefs and behaviours change depending on how social media websites are designed. Theoretically, it contributes by conceptualising the process of variation, selection and retention of knowledge in social media, and proposes a model to study selection through choice-making, while making use of identity, groups and network theories. Further, one of the biggest implications of this research has been to detect how certain elements of website designs affect the transmission of errors and biases. Therefore, if some of the guidelines outlined on this thesis are followed and the designs of websites receive proper attention, the evolution of knowledge in one of the biggest repositories of information for the human kind – social media sites – can be positively affected. Further, methodologically, the study represents one of the most complete examinations performed on social media. Regarding practice, the thesis presents recommendations for practitioners by creating a schematic that can assist them in identifying a more favourable way to set up a website, or at least to detect which features to avoid. Concerning policy, the study reflects on the ethical implications of adopting a frame and raises a number of questions that can serve as the basis of debate among legislators. Overall, it is believed that this research provides a better understanding of the impact that different designs have on the choices of individuals.

CHAPTER 2: LITERATURE REVIEW

Every day people use the internet to communicate, search for entertainment, read the news, and buy and sell products and services, to mention just a few of the main reasons it is used (Correa et al. 2010). The internet has now become an intrinsic component of people's daily lives, and is changing the way they do business, spend leisure time, and most importantly, the way they look for and choose information (Hughes et al. 2012). In addition, the internet is now characterised by the sharing of user-generated content (UGC), defined as diverse types of media content that are both created and utilised by end-users (Kaplan & Haenlein 2010), with social media platforms in particular allowing individuals and communities to produce, alter, share and discuss UGC (Kietzmann et al. 2011).

The present thesis focuses on understanding how the designs of social media sites might affect the way in which individuals share UGC and the effects on identity, groups, and relationships. In order to do this, the literature review firstly describes social media: its definition, classification, and how it differs from face-to-face communication. Secondly, the transmission of UGC is explained through an evolutionary perspective, outlining how content is varied, selected and retained. Thirdly, emphasis is given to the selection of information, which is described through the choice-making process. This is followed, fourthly, by an explanation of how individuals are influenced by groups and relationships when making choices, and three group-biases are described: content, prestige, and conformity. Finally, the fifth section of this chapter presents the theoretical framework and outlines the model, aim, objectives, and research questions of the research.

2.1 SOCIAL MEDIA: DEFINITION, CLASSIFICATION AND FOCUS

This section defines social media and describes the focus of the study regarding three of its building blocks: identity, groups and relationships. Further, the classification of social media is presented with a focus on comparing sites with regards to different degrees of self-presentation (i.e. user profiles) and self-disclosure (i.e. rating scales). Finally, the role of identity is discussed through contrasting offline and online environments.

2.1.1 Definition of Social Media

Social media uses technology, and specifically the internet and digital media, to enable users to collaborate, connect, communicate and interact with one another (Correa et al. 2010; Kaplan & Haenlein 2010; Dabbagh & Kitsantas 2012; Kietzmann et al. 2012; Peters et al. 2013). While a variety of definitions exist, two are most suitable for this study. The first draws from communication science and sociology (Peters et al. 2013) and defines social media as "communication systems that allow their social actors to communicate along dyadic ties" (Peters et al. 2013, p.282). Similarly, the second definition describes social media in terms of a technological platform enabling the transmission of data, and specifically "a group of internet-based applications that build on the ideological and technological foundations of Web 2.0[2], and that allow the creation and exchange of user-generated content" (Kaplan & Haenlein 2010, p.61).

These two definitions highlight four elements that are essential to the understanding of social media: communication; social networks; technology; and UGC. The first two, communication and networks, require the study of social media to take into account information shared within social structures. That this happens within a technological platform means differences between online and face-to-face communication must be taken into account, such as the amount of shared information, (a)synchrony of messages, (a)nonymity of users, and the replacement of spoken language and gestures by text. Finally, unlike other types of media (e.g. print, television), content shared in social media is mainly user-generated, which means that it is created, made publicly available, and used by individuals in the network (Kaplan & Haenlein 2010).

According to the *Organisation for Economic Co-operation and Development*, (OECD 2007), there are three basic characteristics of UGC. First, it must be publicly available online; that is, at least available for a particular group of users on the web, hence excluding emails and instant messages. Second, it should involve some creative effort and some element of originality rather than merely being a copy of existing material. And third, it must be created outside of professional routines and practices, so it should not be

[2] "Web 2.0 is a term that was first used in 2004 to describe a new way in which software developers and end-users started to utilize the World Wide Web; that is, as a platform whereby content and applications are no longer created and published by individuals, but instead are continuously modified by all users in a participatory and collaborative fashion" (Kaplan & Haenlein 2010, p.60-61).

part of a commercial market context. This research will likewise only consider content to be user-generated if it fulfils these three characteristics.

It has been claimed that social media can be understood in terms of seven ‘functional building blocks’: identity; groups; relationships; reputation; presence; conversations; and sharing (Kietzmann et al. 2011; Kietzmann et al. 2012). *Identity* refers to the extent to which users reveal themselves; *groups* are how individuals form communities; *relationships* concerns how people relate to each other; *reputation* involves users knowing the social standing of others; *presence* is knowing if others are available online; *conversations* concerns communicating with other users; and *sharing* involves sending and receiving content (Kietzmann et al. 2012). The present thesis includes elements that encompass all of these blocks. However, a particular emphasis has been placed on identity, groups and relationships. This is because, as seen from the chosen definitions of social media, the communication of individuals within networks is key to the transmission of UGC. However, social media comprises an extensive and varied range of sites, and the term communication is too broad. Thus, the following sections will further explain and refine these terms to narrow the scope of the research more appropriately.

2.1.2 Classification of Social Media

As previously mentioned, social media describes a communication channel that stores and delivers UGC to individuals within social networks. However, the term social media comprises a wide range of site types, set up in very diverse ways, with online environments allowing for various forms of UGC to be shared, using a range of communication processes, and enabling groups to be formed in a number of ways. For this reason, it is relevant to outline the differences and similitudes among sites and to specify at the outset which types of sites are included in this thesis and which are not.

Kaplan and Haenlein (2010) outlined one of the most comprehensive classifications of social media. They proposed two main partitions and two main theories falling into each of those elements. The first division has to do with media research and comprises the theories of social presence and media richness. In an online context, *social presence* is defined as the physical and visual contact that can be achieved by two online users, while *media richness* is the amount of information allowed to be transmitted in a particular time interval (Kaplan & Haenlein 2010). The second partition concerns social processes, encompassing theories of self-presentation and self-disclosure. In the context of social media, *self-presentation* is seen as how users present themselves online, while *self-disclosure* is the revelation of personal information such as opinions, likes and dislikes, consistent with the image users want to present about themselves (Kaplan & Haenlein 2010). Table 2.1 shows the classification described, with examples of pertinent online platforms for each category.

Table 2.1 - Classification of Social Media (Kaplan & Haenlein 2010, p.62)

	MEDIA RESEARCH: social presence & media richness			
SOCIAL PROCESSES: self-presentation & self-disclosure		**Low**	**Medium**	**High**
	High	Blogs	Social networking sites (e.g. Facebook)	Virtual social worlds (e.g. Second Life)
	Low	Collaborative projects (e.g. Wikipedia)	Content communities (e.g. YouTube)	Virtual game worlds (e.g. World of Warcraft)

It should be emphasised that most research on social media has been conducted in relation to only one of the categories shown above – in isolation – and avoids contrasting different types (e.g. Dellarocas 2003; Godes & Mayzlin 2004; Zhao et al. 2008; Thelwall et al. 2011; Wu et al. 2011; Wilkinson & Thelwall 2012; Liu & Park 2015; Park & Nicolau 2015; Jacobsen 2015). Among those who have performed comparisons, a number compare online and offline features (Godes & Mayzlin 2004; Zhao et al. 2008; Berger & Iyengar 2013). Nonetheless, the volume of research comparing social media sites is very small, and most studies examine factors such as user personalities or perceived enjoyment in different social media sites (Hughes et al. 2012; Quan-Haase & Young 2010). Moreover, only two papers were found contrasting platforms by emphasizing the design of different types of social media (Brandtzæg et al. 2009; Li 2017). Hence, some researchers have argued there is a need for more studies focusing on the design of different social media sites to understand better the effects that these have on identity management, self-presentation, self-disclosure, and real-life relationships (Kietzmann et al. 2012).

The first literature gap is therefore the need to investigate how the designs of the different classes of social media sites might enable, affect, or restrain the transmission of information (LG-1). To address this gap, the current research concentrates on the comparison of social media sites with the same social presence and media richness, but varying levels of self-presentation and self-disclosure (see table above). Namely, this thesis studies and compares certain features of *social networking sites* (SNSs), which are characterised by allowing users high levels of self-presentation and self-disclosure, and *content communities* (CCs) which have a lower degree of social processes.

The primary distinction between SNSs and CCs is that the purpose of the former is to connect users by sharing personal information like photos, comments, and thoughts; whereas the objective of the latter class is the distribution of media content such as videos, presentations, or travel advice (Kaplan & Haenlein 2010). Alternatively, it could be said that SNSs rely on relationships whereas CCs are built around topics. However, in the past five years, these two categories have acquired features from one another, making the line that separates them very thin and blurring the boundaries. For instance, YouTube and TripAdvisor (examples of CCs) are now suggesting users sign-in with their Google+ or Facebook accounts (examples of SNSs), thus increasing the amount of self-presentation and self-disclosure of their members. And yet, this process does not appear to have been researched or documented, nor any consequences it may have on the transmission of information. Thus, it can be argued that it is important to have a greater understanding of the impact that different degrees of self-presentation and self-disclosure have on identity management and relationships; and the effects that all of these social dynamics might have on the transmission of UGC.

However, comparing information transmission between two types of social media can be complicated, as the content shared can vary significantly. For this reason, the present chapter offers an in-depth description of both SNSs and CCs, making use of several examples from some of the most well-known sites. Further, it should be emphasised that in the empirical study forming the basis of this thesis, the main features of these categories of social media will be emulated in a *single* site where the same type of information is transmitted.

2.1.3 Identity management in online environments: self-presentation and self-disclosure

When a person is in the presence of others, they will try to obtain information about her, such as her conception of self, socio-economic status, background, trustworthiness, and competences (Goffman 1959). In the same way, when an individual is before others, she will have several reasons for seeking to control the impression she gives, such as making others think highly of her, ensuring harmony in the relationship, or misleading people (Goffman 1959). The *self-concept* has been defined as an individual's thoughts and feelings in reference to herself, and comprises of the *extant-self* (i.e. how she sees herself), the *desired-self* (i.e. how she would like to see herself), and the *presenting-self* (i.e. how she shows herself to others) (Rosenberg 1979). The present research focuses on the last of these, which is intrinsically related to the concept of *identity*, the part of the self by which individuals are known to others (Altheide 2000). It should be noted that identity is not an individual attribute, but instead a product of socialisation (Goffman 1959; Altheide 2000). Therefore, as with communication, identity needs to be studied in a social context.

In online environments, identity can be seen as the extent to which users reveal themselves, or the amount of personal information that social media sites allow being shared (Kietzmann et al. 2012). Some researchers conducting online research have referred to this as 'identity management' (Suler 2002; Suler 2004; Meng & Agarwal 2007) or 'identity construction'(Schau & Gilly 2003; Zhao et al. 2008). For this thesis, identity management corresponds to self-presentation and self-disclosure. Self-presentation – the way users present themselves in cyberspace – is studied through user profiles; self-disclosure – the revelation of personal information such as likes and dislikes – is observed through the ratings that users provide.

Self-presentation in SNSs and CCs: user profiles

Self-presentation can be seen as a component of identity by which an individual tries to make an impression on others (Goffman 1959). This impression is subject to elements that people cannot change termed *personal front* (e.g. gender, age, racial characteristics), components that are under their control called *setting* (e.g. clothing, personal adornments), and a person's *performance* (i.e. actions at a given moment which serve to influence others) (Goffman 1959). An essential feature of the internet is that it gives individuals the opportunity to alter their identity to an extent that would not be possible in face-to-face interaction, allowing them to change even personal front aspects such as their age, gender, and appearance (Suler 2002). What is more, online users can interact with others without making use of a personal front or a setting. Then, by allowing features like disembodiment and anonymity, online environments allow for a new means of identity construction (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008).

Within social media environments, individuals are free to design their physical forms (e.g. avatars, human, animal, hybrids), gender, and any wished symbolic connotations (Schau & Gilly 2003). Nonetheless, as this research focuses only on SNSs and CCs, the way in which these allow individuals to self-present is described through the user profiles that each category allows. SNSs are web-based services that permit users to create a profile, produce lists of other users with whom they want to share information, and allow them to navigate their list of connections to see other people's information (Amichai-Hamburger & Vinitzky 2010). Conversely, in CCs, users are usually not required to create a profile if they only wish to see the available information, although some sites do require very basic profiles when people want to post comments (Kaplan & Haenlein 2010). Figure 2.1 presents the profile of a user in a SNS, while Figure 2.2 shows the profile of the same individual in a CC.

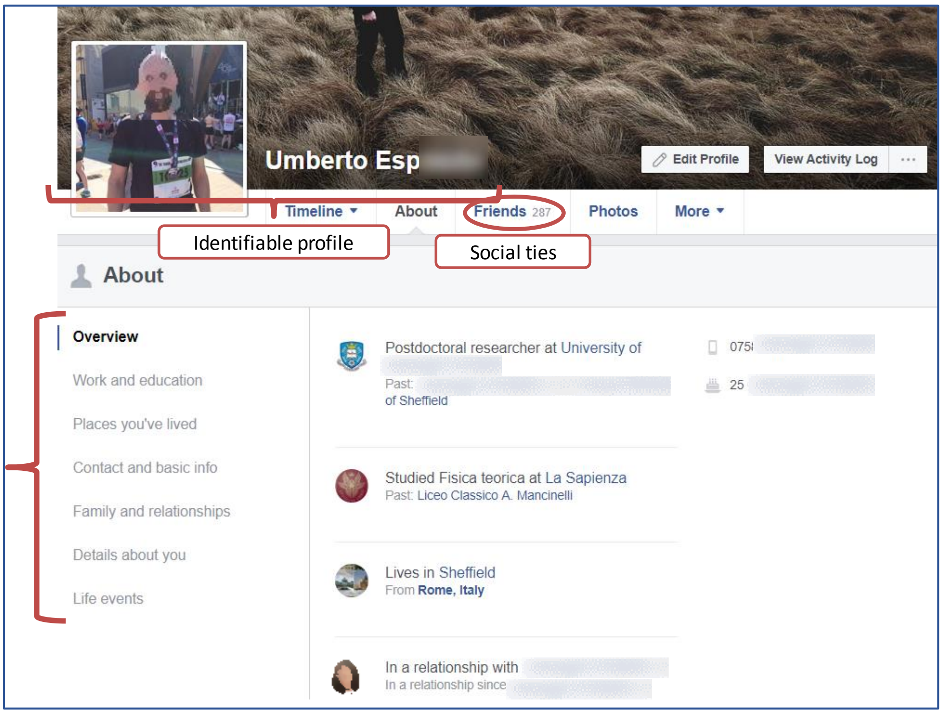


Figure 2.1 - User profile on a SNS, Facebook

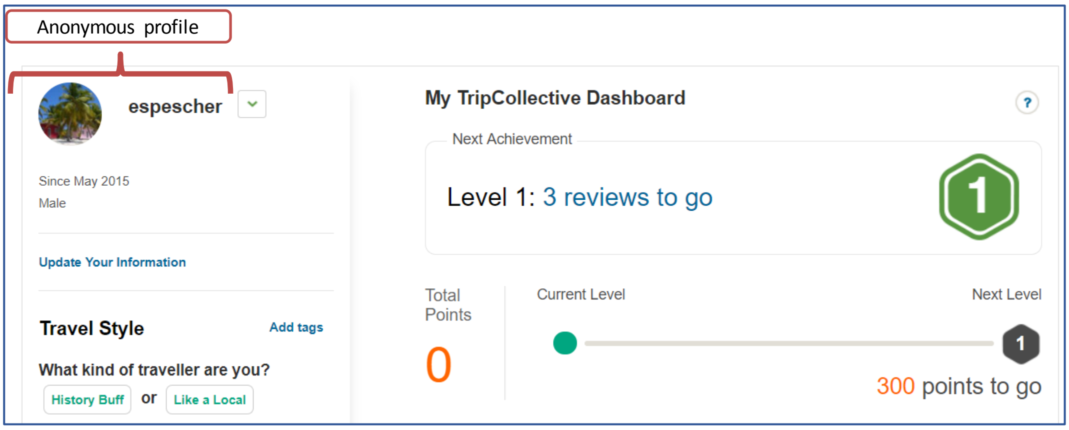


Figure 2.2 - User profile on a CC, TripAdvisor

As can be seen from the images presented above and as explained in Table 2.1, SNSs allow for and are characterised by a much higher degree of self-presentation. For instance, Facebook (Figure 2.1) began by requiring users to sign up with a valid university email (Loomer 2012), thus linking users to their real identities. Moreover, it was meant to act as 'a facebook' where students would display a real picture and write something about themselves. Also, for users to become 'friends' there needs to be a mutual agreement, which arguably makes users more likely to use their real identities in order to be recognised by others. Users are typically encouraged to provide as much information about themselves as possible, such as current and previous jobs and education, relationship status, list of family members within the SNS, and significant life events. Similar levels of self-presentation can be found in SNSs such as Instagram, Google+, LinkedIn, and Snapchat.

Conversely, CCs do not typically require users to display their real names, and have significantly fewer fields displaying personal information. For instance, the TripAdvisor account in Figure 2.2 does not display any name, as this is not a requirement. Moreover, there is nothing in place for users to include information about themselves. In contrast with Facebook, TripAdvisor does not allow for 'friends' or 'followers', presumably emphasising the content being shared rather than social bonds. However, it should be said that a few CCs, like YouTube and Quora, do allow users to have 'followers', but these are 'directed' so only one party has to agree to form a connection (i.e. the following need not be reciprocated). To sum up, regarding user profiles, the most distinctive characteristic that differentiates CCs from SNSs is that the former allows for anonymity, mainly because the emphasis is on the content being shared. In contrast, in the latter, users are defined as '*nonymous*'[3] (i.e. non-anonymous) because in these sites relationships are anchored to offline environments, which means users are meant to be known by others to a certain extent (Zhao et al. 2008).

The role of anonymity in online environments has been widely studied (e.g. Lea & Spears 1991; Reicher et al. 1995; Postmes et al. 2001). Further, there seem to be two schools of thought as to why individuals behave differently when being anonymous as opposed to identifiable. Some research is based on *deindividuation theory*, which argues that there is only one concept of 'self' that is reduced to group norms when an individual is within a network (Le Bon 1896; Zimbardo 1969). Deindividuation theory is mostly used as a justification for anti-normative behaviour. Researchers making use of this theory claim that users lose their 'self' to the group, a situation that aggravates further when individuals are anonymous and communicate through a computer, as opposed to face-to-face (Lu & Bol 2007; Brandtzæg et al. 2009). On the other hand, other researchers rely on the *social identity model of deindividuation effects* (SIDE), which claims that individuals do not lose themselves in the network because the self-concept adapts to diverse contexts and situations (Lea & Spears 1991; Reicher et al. 1995; Postmes et al. 2001). Researchers justifying interactions through SIDE recognise that in some situations groups can behave negatively, but also argue that anonymity can generate altruistic acts among group members, although this would depend on the norms of each group (Howard et al. 2010; Whittaker & Kowalski 2015; Mishna et al. 2009; Douglas & McGarty 2001). The present research follows a SIDE perspective as it is in line with the identity theory used in this thesis, which assumes both that individuals vary their 'self' depending on the circumstances (Goffman 1959; 1963).

Unlike anonymity, "identity construction in a nonymous online environment has not been well studied" (Zhao et al. 2008, p.1818). That is, there is a lack of research exploring the way in which individuals manage their identities when they are known by other members of the network (LG-2). In one of the few studies to date, Zhao et al. (2008) found that when users are identifiable, they do not change their personalities but rather enhance certain aspects of these. In identity jargon, this could be described as individuals highlighting their 'desired-self'.

[3] Note that the term 'identifiable' will generally be used as a synonym for 'nonymous' throughout.

Finally, one last term to be discussed within identity is what Goffman (1959, p.60) called ‘embarrassments’, which occur when an individual is caught performing in a manner that is not consistent with their ‘official projection’. In this regard, Kietzmann et al. (2012) have emphasised the need for a greater understanding of how embarrassments unfold: how people react to incongruent information about others, how these affect real-life interactions, and the impact they can have in the design of social media platforms (LG-3). To address these last two gaps, LG-2 and LG-3, the present study compares an anonymous with a non-anonymous environment and, with identifiable users, it investigates the effects that online interactions can have in face-to-face situations and vice-versa.

Self-disclosure in SNSs and CCs: rating scales

As described above, identity comprises self-presentation and self-disclosure, and SNSs are characterised by higher levels of these social processes in comparison to CCs. Self-disclosure is defined as the revelation of personal information aligned with the image online users wish to give about themselves (Kaplan & Haenlein 2010). Thus, it should be noted that individuals’ disclosure is aligned and intrinsically related to their self-presentation. Furthermore, self-disclosure may involve information about the users’ emotions, attitudes, feelings, thoughts, likes and dislikes (Moon 2000; Kaplan & Haenlein 2010). For this study, self-disclosure is studied through the likes and dislikes that users expressed through the available rating scales.

Nowadays, mechanisms for user ratings are available on almost every website. This is happening due to the proliferation of social participation and co-creation within social media (Riedl et al. 2013). Rating within an online community reflects a measure of the quality of an idea, and rating scales are used to assess almost every imaginable category, such as videos, movies, consumer electronics, travel services, teachers, coding, and books (Riedl et al. 2010). Therefore, ratings also serve as online feedback mechanisms that disseminate word-of-mouth information about products, experiences and services within networks (Dellarocas 2003). Hence, rating scales act as a medium for users to disclose information about themselves (i.e. their likes and dislikes), but also as feedback mechanisms because individuals are encouraged to use them to share information about products and services (e.g. TripAdvisor and Amazon).

Rating scales in online environments are under-studied, and there are no clear guidelines regarding their designs, and how these impact on the effectiveness and quality of ratings (Riedl et al. 2010; Riedl et al. 2013). Although there is a significant amount of research focusing on consumer ratings (e.g. Jiang et al. 2015; Park & Nicolau 2015; Liu & Park 2015; Gopinath et al. 2014; Tsao et al. 2015), the *status quo* of the available rating scales on sites is rarely questioned. Further, from the limited number of studies that question and compare different rating scales, the recommendations seem to be contradictory. On the one hand, academics tend to suggest that longer scales are needed (Riedl et al. 2010; Riedl et al. 2013); on the other hand, practitioners prefer using shorter rating scales (YouTube 2009; Ciancutti 2011). Hence, the lack of consensus regarding the design of rating scales and the absence of studies that investigate how these affect the transmission of information can be considered a gap concerning both literature and practice (LG-4).

In respect of existing rating scales, SNSs are typically characterized by using a one-point rating, as can be seen in Table 2.2. Conversely, most CCs use more granular ratings which – as seen in Table 2.3 – vary between one and ten-point scales, although the majority uses a scale from one to five. Both tables present the rating scales that the sites had in 2015 and the ones they have now, in 2018. It was decided to leave the 'old' and 'new' scales, so the rapid changes that both SNSs and CCs are having can be appreciated. Further, as will be explained in the following paragraphs, most of these alterations seem to be made by trial and error with their effects apparently largely unknown a priori.

Table 2.2 – Rating scales of SNSs (2015-2018)

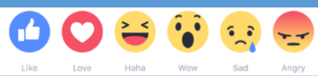
Social Networking Site	Name of button (2015 → 2018)	Rating scale (Aug/2015)	Rating Scale (Apr/2018)
Facebook	Like → Reactions		
LinkedIn	Like		
Google+	+1		
Instagram	Like		
Tumblr	Like		
Pinterest	Like & Pin-it → Save	 & 	

Table 2.3 – Rating scales of CCs (2015-2018)

Content Community	Rating-scale type	Rating scale (Aug/2015)	Rating Scale (Apr/2018)
YouTube	Likert [4] → Dichotomous	4.6	857 30
Twitter[5]	One-point	Favorite	Like
Quora	Dichotomous	Upvote 1.2k Downvote	Upvote 1.2k Downvote
Stack Overflow	Dichotomous & one-point	187 54	187 54
TripAdvisor	Likert (1-5) [6]		4.5
Booking.com	Grid (1-10)	9.1	8.1 Very Good
Couchsurfing	Likert → Dichotomous	Positive (74) Neutral (1) Negative (0)	Would Stay Again/Positive (8) Would Not Stay Again/Neutral/Negative (0)
Yelp.com	Likert (1-5)		
Amazon	Likert (1-5)		
Netflix	Likert → Dichotomous		
Google reviews	Likert (1-5)		4.8

Even though SNSs allow for higher levels of self-disclosure, they also have smaller scales in place, arguably reducing the amount of feedback that users can give to others. Interestingly, the site that has been 'blamed' for the widespread use of the 'like' and 'love' buttons was the first to adopt a wider scale, thus offering more options on how to respond to the posts of other users. As of today, Facebook is the leading SNS with 2,167 million active users worldwide (Statista 2018a). Moreover, it is thought that more than a third of the UK and US populations access Facebook every day (Sedghi 2014). Regarding its rating scale, Facebook had been widely criticised for not allowing users to 'dislike' posts, which made the website a very attractive site for politicians and marketing campaigns (Fraser & Dutta 2008). Therefore, due to the insistence of users and the criticism from different parties, the site introduced a range of 'reactions' in February 2016 (Krug 2016;

[4] YouTube's likert scale was changed for a dichotomous one in 2010, but it was decided to include this image for reasons explained below.

[5] Although some consider Twitter to be a SNS (e.g. Statista 2018a), the official stand of the company is that they are not. Rather, *"Twitter is for news. Twitter is for content. Twitter is for information"* (Perez 2010; McCracken 2016).

[6] Note that some of these websites (e.g. TripAdvisor, Booking.com, Yelp, Amazon and Google reviews) have a two-step rating: 1) people write reviews about products and services, accompanied by a rating using a likert scale; and 2) other users rate the written reviews with a one-point rating 'helpful/useful' button.

Teehan 2016). Consequently, after twelve years of only allowing a single rating button, the site now gives its users the option of higher levels of self-disclosure.

Conversely, CCs have been known for having more extensive rating scales but, as seen in Table 2.3, the sites that have re-designed them are increasingly opting for offering users fewer options to give feedback. An example of this is YouTube, the biggest CC, where its 1,500 million users share millions of hours of video every day, available in 61 languages (YouTube 2015; Statista 2018). Until 2010, they had a 5-point likert scale, which then became dichotomous (i.e. ‘thumbs up, thumbs down’). This change, they disclosed, was because users would sometimes rate videos when they did not like them, but overwhelmingly rated them when they really enjoyed them, as seen in Figure 2.3 (YouTube 2009). Thus, the video-sharing site asked its users what would they find more useful: a ‘like/dislike’ scale or a ‘favourite’ button, and the dichotomous scale was adopted (YouTube 2009).

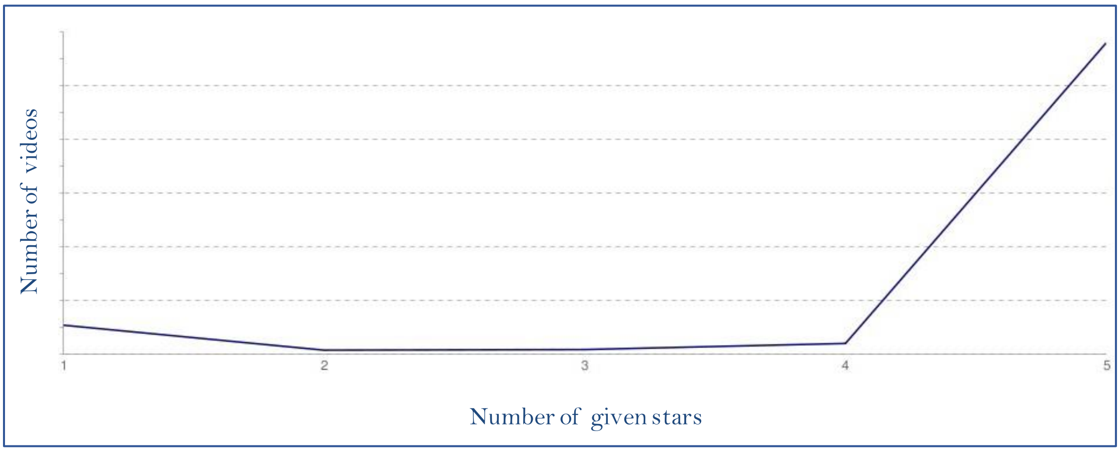


Figure 2.3 – YouTube's likert 1-to-5 rating chart (YouTube Official Blog, 2009)

Citing another example, a product engineer in charge of the design of Netflix’s rating scale described a real-life experiment that they made with their 5-point likert scale, during 2011 (Ciancutti 2011). For a limited time, half of the users were allowed to rate using whole-numbers (i.e. complete stars), while the other half could give whole or half-points (i.e. half starts). This experiment was done because it was considered the software could have given ‘better’ movie suggestions if users were allowed to give more fine-grained ratings that more accurately described their opinions. However, the result was that Netflix members who could rate using half-stars gave significantly fewer ratings, and thus the site continued using a whole-point scale (Ciancutti 2011). It should be noted that in mid-2017 Netflix subsequently changed its feedback mechanism for a dichotomous scale (i.e. 'thumbs up, thumbs down'), and has since received criticism from its users, who find the scale too restrictive and the predictions based on ratings inaccurate (e.g Smith 2017).

From the examples of SNSs and CCs discussed above, it appears that when users are given fewer rating options they feel restricted (e.g. Facebook and Netflix). However, in some situations when they are given more nuanced options to express feedback, they do not use them (e.g. YouTube, Figure 2.3). Therefore, it looks as if the above-mentioned social media sites were designing their scales based on trial and error and from the feedback they received from their users. This may be because guidelines on the usefulness of different rating scales are scarce and contradictory. On the one hand, practitioners in charge of the design of rating scales argue – based on the millions of ratings from their users – that scales should be short, as users rate less when given more complex scales, on top of not using all of the available options (YouTube 2009; Ciancutti 2011). On the other hand, scholars such as Riedl et al. (2010) argue that shorter rating scales such as dichotomous and single-point, force respondents into making a choice regarding the quality of an idea or post. However, the thoughts and emotions that take place during this decision-making process are not likely to be dichotomous, which can cause users to fail in the process of judging between alternatives (Riedl et al. 2010). Consequently, these scholars suggest users may experience high decisional stress. Further, they claim that simple rating scales do not produce valid rankings, as these are obtained with longer multi-attribute scales, which not only produce significantly better ratings but also lead to more favourable attitudes towards websites (Riedl et al. 2010; Riedl et al. 2013).

To sum up, as outlined in the fourth gap (LG-4) there is a theoretical and practical need to deepen the understanding of rating scales. Moreover, given that there is a tendency of CCs to increase users' self-presentation by promoting the idea of signing-in with their Facebook or Google+ accounts, it could be claimed that the study of ratings should be done within the context of users' identities.

2.2 TRANSMISSION OF USER-GENERATED CONTENT IN SOCIAL MEDIA: AN EVOLUTIONARY PERSPECTIVE

Social media was defined in terms of four elements: communication, UGC, technology and social networks (Kaplan & Haenlein 2010; Peters et al. 2013). From a communications perspective, it was seen as a medium for storing and delivering information. However, unlike other mass media, the information shared in SNSs and CCs is created and acquired by users, and has thus been named UGC. Further, technology plays a role in the transmission of UGC because it happens through computers and mobile devices, allowing anonymity of users, asynchrony of messages, and text and emoticons replacing spoken language and gestures. Lastly, UGC is publicly available to individuals within a network, which is also a particular characteristic of communication in social media.

This study uses an evolutionary perspective to make sense of the communication of UGC among networked individuals, through the use technological devices. Moreover, using an evolutionary perspective helps to understand better how the different designs of social media sites (i.e. SNSs and CCs) enable, restrain or affect the transmission of information. The present section starts by defining evolution and by explaining what is evolving. Next, it examines how knowledge is (en)coded into UGC, and vice-versa. Further, the process of knowledge interpretation is explored, with an emphasis on how transmission errors might take place. Finally, all the previously outlined concepts are combined in a schematic summarising conceptually how knowledge evolution takes place in social media.

2.2.1 Evolution: what is it, and what is actually evolving?

The concept of *evolution*, as described by Darwin (1859), assumed that – whatever was evolving – was firstly varied, struggled for existence, went through natural selection and, if successful, was then inherited with or without modification. These concepts, applied nowadays to a number of disciplines, are referred to as the mechanisms of *variation, selection* and *retention* (VSR), first described by Campbell[7] (1960) and later adopted by

[7] Campbell originally used the term *blind-variation-and-selective-retention (BVSR)* to describe the creative knowledge process.

organisational researchers including Weick (1969), Aldrich & Ruef (2006), and Breslin (2011). It should be emphasised that this research assumes a Generalised Darwinist position which allows generalising the VSR mechanisms described by Darwin to domains outside biology and in fields such as of sociocultural evolution (Breslin 2010; Breslin 2011). This position states that Darwinian evolution will occur "as long as there is a population of replicating entities that makes imperfect copies of themselves, and not all of these entities have the potential to survive" (Hodgson 2005, p.900). These replicating entities, named *replicators*, are identified by the *interactors*, which are their visible manifestations in the world of culture (Dawkins 1976). In these terms, *variation* implies that the replicators are sufficiently different from each other in order to be distinguished and later selected; the *selection* mechanism supposes that the units that evolve go through a competition in order to be chosen; and *retention* assumes that the selected entity can be identified to be kept by those who have selected it, or can be inherited either intact or with modification (Mesoudi 2011).

The first scholar to describe the replicators or "social units" was Cloak, affirming that culture was "acquired in tiny, unrelated snippets, which are specific interneural instructions culturally transmitted from generation to generation" (1975, pp.167–8). A year later, Dawkins named them "memes", arguing that these entities were faithfully replicated in a discrete manner, like genes (Dawkins 1976, p.192). Likewise, Distin uses of the term "memes" (2011, p.231) but predominantly refers to (discrete) "cultural information" (2011, p.11), although she also uses "ideas" (2006, p.90) and "information" (2006, p.92). Conversely, other researchers such as Richerson & Boyd have questioned the idea of Darwinian replicating entities, and instead refer to them as (information) "variants", although they also use the (more common) terms "idea, skill, belief, attitude and value" (2005, p.63). Other authors, like Mesoudi, have stayed on the margin of the discrete versus continuous debate but still affirm that there must be some sort of replicators, and describe them as (information) "traits" (2011, p.55). Further, Breslin has referred to "ideas" as "units of evolution" (Breslin & Jones 2014, p.435) but also defines them as "routines" when studying knowledge-in-practice (Breslin 2016, p.49).

For this research, and in the context of social media, the author will argue that the dualism of beliefs/UGC evolves through the mechanisms of variation, selection, and retention. This content is first conceived as the knowledge (i.e. beliefs) that social media users possess, which is then translated into digital information (e.g. text, images, videos, emoticons). That is, the beliefs that individuals hold are the replicators, and the UGC they post and browse online is the interactor. Therefore, knowledge is evolving in the form of beliefs, but to be understood and acquired by others, it needs to be coded into discrete entities of information (i.e. content). However, once online users have acquired the UGC and this is decoded again into beliefs, it is not possible to determine if it is still discrete or continuous, as this is something left to neuroscientists to decide (Mesoudi 2011). In order to explain this process further, the duality of beliefs-UGC will be defined, as well as the interpretation between one and the other.

Nevertheless, before moving into a deeper understanding of knowledge and information, it should be briefly mentioned why these terms were chosen as opposed to 'memes', which might sound appealing given that this research takes place in social media. As previously explained, the word 'meme' was coined by Dawkins (1976), so it resembled the biological term 'gene'. Therefore, by using 'memes', the research would hence adopt a Universal Darwinism posture, which analyses sociocultural aspects in the same way that biology does (Breslin 2010). Rather, this research goes beyond biology and abstractly studies the evolutionary mechanisms, aligned with the previously outlined philosophy of Generalised Darwinism. Consequently, this research does not imply that biology and culture, or genes and memes, behave in the same manner. Moreover, 'meme' carries with it the connotation that within social media "the most common meme is an image of a person or animal with a funny or witty caption" (Beal 2016). Thus, the term 'meme' has been overexploited and misapplied by the users of the internet, which has caused a deterioration in the credibility of 'memetics' (Distin 2011). Consequently the term is avoided in this thesis.

2.2.2 Duality and interpretation between beliefs (i.e. knowledge) and UGC (i.e. information)

Could a person claim to know how to ride a bike if they have only read about the bicycle's velocity over the angle of imbalance? Probably not (Polanyi 1968). Knowledge is different from information. There are many ways to define and describe knowledge, a term that conveys multi-layered meanings (Nonaka 1994). Nonetheless, two definitions were considered for this study. First, knowledge has been defined as "a justified true belief" (Nonaka 1994, p.15), a definition that is relevant because it identifies the 'entity' that is evolving, the so-called 'replicator' described in the preceding paragraphs. Further, knowledge has also been described as "an ongoing social accomplishment" (Orlikowski 2002, p.252), which highlights the importance of the social element; the evolution of knowledge cannot happen within a single individual.

Likewise, information is not self-contained, as it requires a context to acquire meaning (Loasby 2002). Further, information can be described as the way knowledge is translated, the code it uses to take form (Barthelme et al. 1998). In other words, "information is a flow of messages, while knowledge is created and organised by the very flow of information, anchored on the commitment and beliefs of its holder" (Nonaka 1994, p.15). Furthermore, one of the primary conditions of knowledge transmission is that information must be communicated to a receiver who can understand it and respond appropriately (Distin 2011). Therefore, the transmission of information alone would not provide the conditions for knowledge evolution; for this to happen, interaction needs to take place and therefore knowledge is not only transferred but is also transformed by the replication and variation that happens through communication (Dobson et al. 2013).

In this regard, the way knowledge is transmitted, interpreted, and transformed has long been studied by researchers (Polanyi 1968; Polanyi 1967; Polanyi 1998; Nonaka 1994; Orlikowski 2002). Polanyi (1968; 1998) was one of the first scholars to detect that people could know more than they could put into words, and applied the terms 'tacit' and 'explicit' to the study of knowledge. Afterwards, a number of scholars applied these concepts to the study of knowledge transmission within groups and organisations. *Explicit knowledge* was defined as codified knowledge which is "transmittable in formal, systematic language", while *tacit knowledge* was conceived as hard to explain or communicate and is "deeply rooted in action, commitment and involvement in a specific context" (Nonaka 1994, p.16). This view of knowledge seemed to typecast explicit knowledge as "digital […] captured in libraries, archives and databases" (Nonaka 1994, p.17) while tacit knowledge was "constituted and reconstituted in everyday practice" (Orlikowski 2002, p.252). Although these definitions describe some of the characteristics of tacit and explicit knowledge, they also seem to simplify them to the point where everything that is written is considered explicit, while all that involves performing an activity is classified as tacit. However, it should be noted that these categorisations of tacit and explicit knowledge are far from how these definitions were initially conceived.

Polanyi (1967; 1968) advanced the idea that the use of languages was pure tacit knowledge, even if this transmission of knowledge was only conducted by writing. With the example of writing a letter, he illustrated the process of conveying and interpreting meaning. To begin with, the person writing a letter would first require an intelligent understanding of events (*sense-reading*) which would then have to be put into prose (*sense-giving*), and the person receiving the text would have to interpret the written composition to reproduce the original meaning (*sense-reading*). Finally, the person reading the letter would not remember all the words or sentences that composed the letter, but rather the meaning it conveyed. Further, he added that "all knowledge falls into one of these two classes: it is either tacit or rooted in tacit knowledge" (Polanyi 1967, p.314).

Likewise, it could be argued that a similar process takes place in social media. Initially, the knowledge users have (i.e. beliefs) is translated into information (i.e. UGC in the form of text, videos, emoticons, etc.) by the act of sense-giving. Afterwards, once the UGC is available online, other users might acquire it, by decoding this information back into knowledge (i.e. sense-reading). There are two points that should be highlighted. First, the person reading the information posted online must understand the code (e.g. language and signs) to extract its meaning. Second, even if the person who reads the UGC understands the language, they will possibly give a different meaning to it; that is, they may have a different interpretation. This can happen because people rely on their prior understanding of experiences to give-sense, while communicating (Polanyi 1967). Different interpretations explain why beliefs are 'inherited' with modification (Mesoudi 2011). Figure 2.4 shows how knowledge (i.e. beliefs) are translated into information (i.e. UGC) through the interpretation process explained by Polanyi (1967). Therefore, the figure below also displays how replicators are converted to interactors, within online environments.

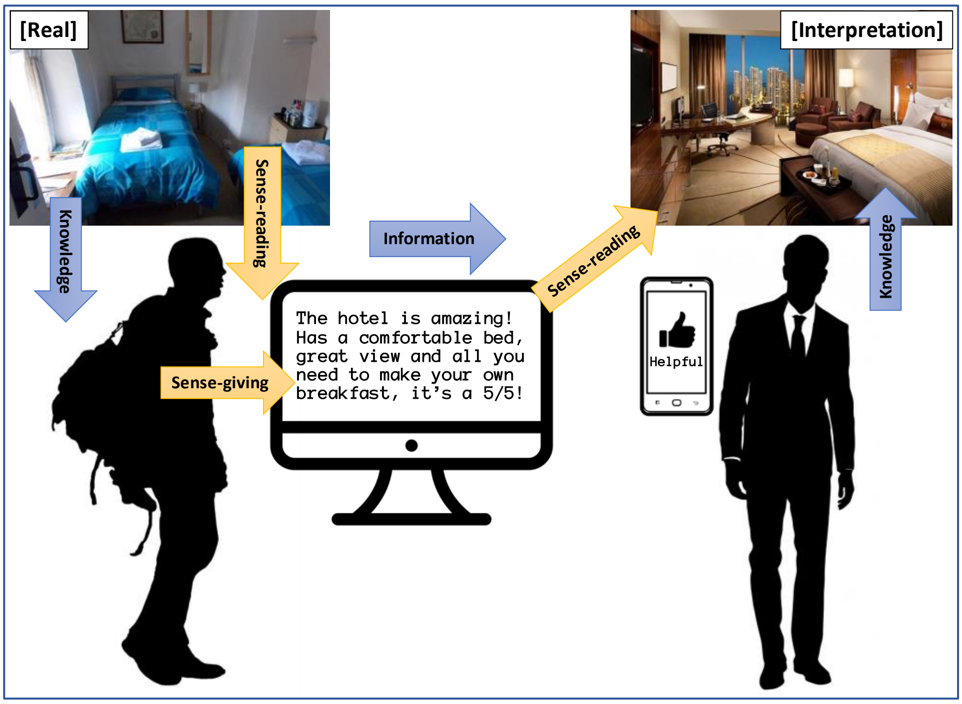


Figure 2.4 - Knowledge, information, sense-reading, and sense-giving in social media

2.2.3 Knowledge evolution in social networking sites and content communities

The way in which knowledge evolution is thought of taking place in social media is shown in Figure 2.5. As can be viewed, variation can be seen as the pool of information that online users generate within the SNSs or CCs, where they communicate their beliefs, through text, audio, video, symbols, emoticons, and so forth. The selection part takes place when an individual within the network chooses a particular piece of information. As outlined previously, this is when the information is decoded into knowledge, and the belief is replicated. Finally, the retention mechanism takes place when the person assimilates the belief, which may differ from the original due to interpretation (Figure 2.4). The way in which the VSR mechanisms can be observed will vary with the different designs that SNSs and CCs have in place. However, the diagram below outlines a general way in which they can be detected. It should be highlighted that Figure 2.5 is a first step to envision how the VSR process occurs in social media. However, Chapters 8 and 9 will reflect and amend this original proposition (e.g. see Figure 8.4).

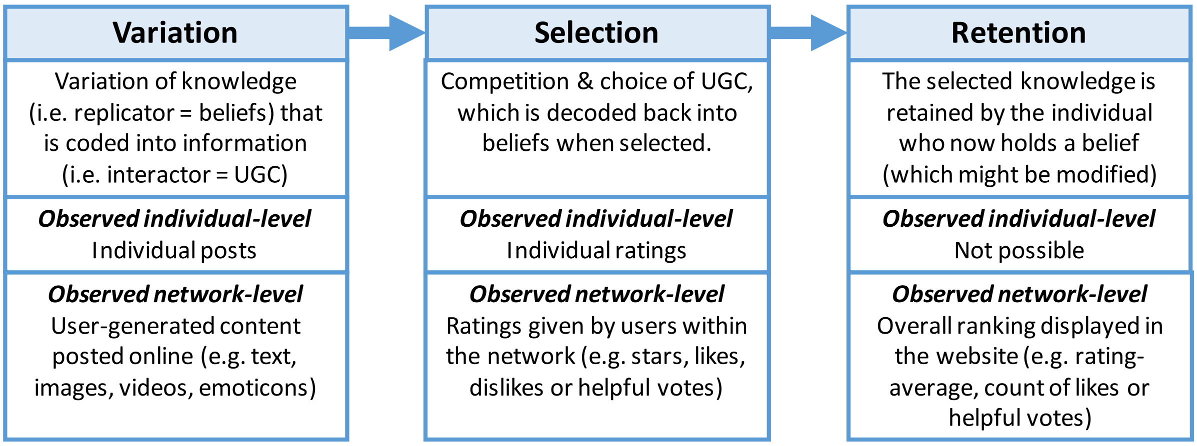


Figure 2.5 – Envisioned VSR process in social media

<u>*Individual level*</u>

Variation happens when online users share their knowledge about a particular topic, product, service, or experience. There is no way of seeing the replicating entities (i.e. beliefs) other than through the corresponding interactors (i.e. UGC) which are subject to the allowed media type(s) and buttons that different websites have in place. For instance, on YouTube the variation can be observed through each posted video, on TripAdvisor through reviews, and on Facebook through posts. However, selection and retention are not as straightforward. For instance, when people read a post or watch a video, they might acquire knowledge without even being aware of it, i.e. unconsciously (e.g. Campbell 1960; Polanyi 1967). It would then be impossible to test every online user and decide whether they have selected or retained knowledge. The selection mechanism is the hardest to detect, especially when talking about beliefs because "a valid question might be 'how do you know it was selected?' and the answer often is, 'because it is there'. It is there because it was selected, it was selected because it is there" (Weick 1969, p.56). Then, the only way in which the researcher can know if the belief was selected is if online users 'claim' in some way that they have agreed with the information posted on the SNS or CC. This, as noted before, varies depending on the design that different websites have in place, but will mainly be seen through the rating scales (see Table 2.2 and Table 2.3 for some examples). Finally, it is believed that it not possible to observe retention at the individual-level in an online environment given that it is not possible to test if each user now holds the specific belief. Thereafter, it can be inferred that if a user has 'claimed' to agree with certain content, s(he) now retains the original belief, with or without modification.

Network level

The variation mechanism is relatively easy to detect within the network: as long as there are differentiated texts, images, emoticons, videos, and reviews, there is variation. Moreover, as with the individual-level, the selection occurs through the ratings of users which can be seen through the available rating buttons, such as stars, up-votes, down-votes, and helpful votes, among others. Finally, retention can be conceived as the opposite of variation, the latter implying novelty (Weick 1969). Hence, retention can be detected when most of the network agrees with a belief. For instance, through the rating-averages, rankings, and sum of 'likes', 'up-votes', or 'love' votes. Figure 2.6 presents an example of how the VSR mechanisms can be observed at the network level in the CC of StackOverflow[8]. As can be seen, the variation would be all the answers posted by users, selection would be the individual ratings (i.e. up-votes), and retention would be the ranking of questions and answers, obtained through the 'wisdom' of all network members.

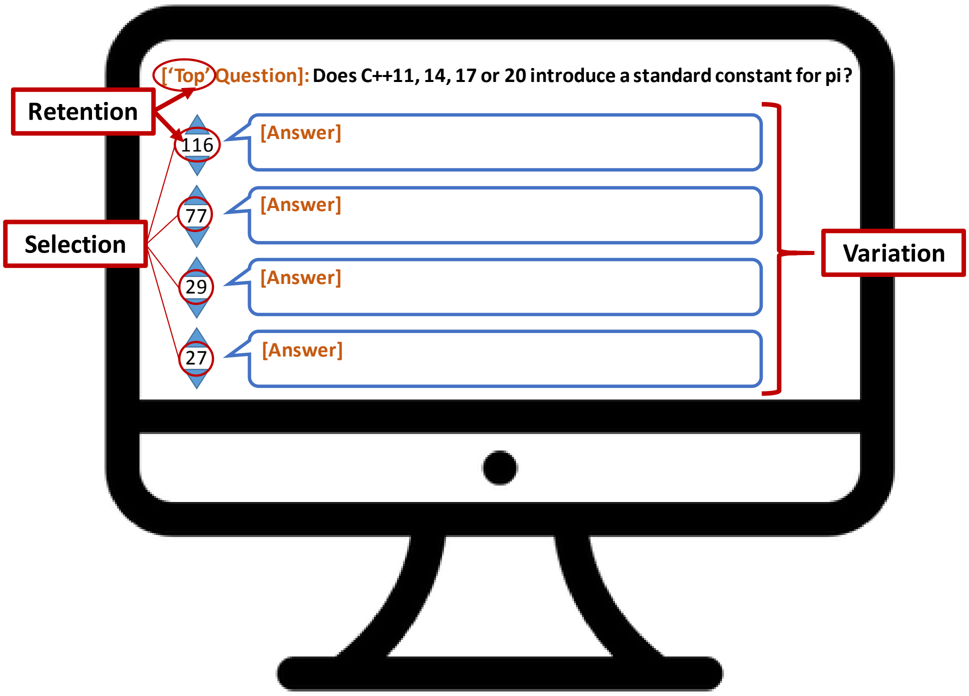


Figure 2.6 - VSR mechanisms shown at network-level (example from StackOverflow)

A gap within the literature is that – from an evolutionary viewpoint – the transmission of knowledge has not received sufficient attention, despite proving to be significantly useful

[8] Note that this figure is based on a real-life example which was take from the answers of a 'top question' on 19/04/18 (see https://tinyurl.com/ydy2jsgc).

in explaining the diffusion processes at the individual and network levels (LG-5). As Richerson and Boyd (2005, p.123) argue, "Darwinian analysis reveals a mass of largely unexplored questions surrounding the psychology of knowledge transmission and the biases that affect what we learn from others. Small, dull effects at the individual level are the stuff of powerful forces of evolution at the level of populations. Understanding precisely how individuals deploy their kit of imitation heuristics is necessary to understand the rates and direction of [knowledge][9] evolution, and work on the problem has hardly begun". It should be noted that the outlined gap gives more importance to 'imitation heuristics' – that is, selection – as this is responsible for individuals acquiring beliefs that are later shared by the network. For this reason, the present study adopts an evolutionary approach to the study of online environments, and focuses on the selection mechanism, explored through choice-making.

Finally, it should be mentioned that the lack of Darwinian concepts in the study of knowledge transmission is even more evident in studies concerning online environments. In this field, there are very few scholars applying the concept of knowledge evolution (e.g. Chen & Liang 2011; Kump et al. 2013; Jiang et al. 2014). Moreover, of these, only Chen & Liang (2011) mention the VSR mechanisms. This highlights the need for more research that understands how variation, selection, and retention take place in online environments. Therefore, the current study attempts to contribute to the field by describing the VSR mechanisms of knowledge evolution within social media, emphasising in particular selection, i.e. how and from whom online users select information. Further, this has been studied through the comparison of the characteristic user profiles and rating systems of SNSs and CCs.

2.3 EXAMINING THE SELECTION MECHANISM THROUGH CHOICE-MAKING

The transmission of beliefs-UGC in social media is studied through the VSR mechanisms. The primary interest of this thesis is on selection, which can be inferred from the ratings of online users. In this respect, scholars studying ratings have claimed that people go through a choice-making process similar to that of answering surveys, as their answers

[9] The term 'cultural' has been replaced for 'knowledge', as the first is considered too broad for this study.

are subject to the available response scale. Therefore, the present section begins by outlining the choice-making process of survey response. Afterwards, the section deepens the explanation of choices by describing how the judgement of users is affected by the manner in which decision-problems are framed. Further, heuristics and biases are explained in the context of frames, judgement and choice. Finally, the section ends with the three group-biases that were chosen for this study: content, prestige and conformity.

2.3.1 Rating: a choice-making process

Riedl et al. (2010; 2013) studied the action of performing a rating in a similar manner as responding to a survey, provided that users are presented with content and then given a limited scale with which to assess or respond to it. This thesis has studied the selection of information, inferred from the ratings performed online, in an analogous manner. Therefore, a rating can be seen as the product of choice, which is the outcome of a process that involves assessment and judgement and requires the evaluation of different options and making a decision about which one to report (Hastie & Dawes 2010). Moreover, to perform a rating, users first need to understand the content, then make a judgement about its quality and finally give a rating that corresponds to the available rating scale (Riedl et al. 2010). Figure 2.7 shows a graphic representation of how Tourangeau et al. (2000) have envisioned the process of responding to a survey.

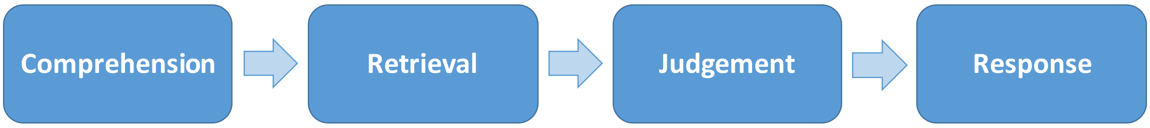


Figure 2.7 - Choice-making process when responding to surveys (Tourangeau et al. 2000, p.8)

The process depicted in Figure 2.7 involves four steps. 1) *Comprehension* comprises understanding the content, making a logical representation of it, and linking the key terms to relevant notions. 2) *Retrieval* is to recall the significant information or memories, and to replace any missing details. 3) *Judgement* consists of assessing the completeness and significance of the recalled information, drawing inferences based on accessibility, integrating the material, and making an estimate of the answer based on partial retrieval. 4) Finally, the *response* involves mapping the judgement into one of the options on the scale (Tourangeau et al. 2000). This framework and, specifically judgement, is further explained by making use of Tversky and Kahneman's (1979; 1981) choice-making theory, explained below.

2.3.2 Prospect theory: choice, judgement & frames

There are several theories that deal with choices and could therefore be applied to ratings within social media, such as the adaptive toolbox decision-maker framework (Gigerenzer 2000), the role of emotion in the choice-making process (Bettman et al. 1998), the dual-process of thinking: system 1 and system 2 (Kahneman 2012), among others. However, the framework considered most appropriate for this topic is *prospect theory* (Kahneman & Tversky 1979). This is because: 1) it is based on the concept of bounded rationality, 2) outlines the concept of frames, and 3) originated the study of heuristics and biases.

During the 1970s, Tversky and Kahneman transformed the idea of the judgement of individuals (Gilovich et al. 2002). Their experiments showed that when individuals were presented with choices regarding uncertainty, risk, or money, they would not follow the then established e*xpected utility model*, which stated that they would try to maximise the gains (e.g. Von Neumann & Morgenstern 1957). Instead, individuals seemed to make choices according to what they called 'prospect theory', which stated that perception was referent dependent in the way that the perceived characteristics of an object or situation differed from those of the same situation when shown (i.e. 'framed') differently (Kahneman & Tversky 1979; Tversky & Kahneman 1981; Kahneman 2003). Prospect theory had its roots in *bounded rationality*, which argued that individuals' choices are subject to internal and external constraints such as their cognitive limitations, the way in which problems are formulated, and the timescale within which they need to make a choice (Simon 1955; Simon 1979). Likewise, the framework outlined by Tversky and Kahneman established that, when making a choice, most people would depend on heuristic principles (i.e. rules of thumb) which would reduce the difficulty of tasks, but at the same time would lead to critical systematic errors, or *biases* (Kahneman 2003).

Tversky & Kahneman's (1981) experiments consisted of a *decision problem*, defined by the options presented to an individual, the possible consequences of her acts, and the likelihood of those outcomes. Moreover, the *frame* adopted by the decision-maker was controlled both by the formulation of the problem and by the person's habits and personal characteristics. Regarding framing, most people seem to commit at least one of the following irrationalities, even when having a definite preference (Tversky & Kahneman 1981, pp.457–8):

1. They would choose a different option if the same question was framed in a different way.
2. They would not recognise alternative frames and the way they affected expected utilities.
3. They would believe their choices were always the same, aligned to their original preference.
4. When they realise their choices were being inconsistent, they did not seem able to resolve it.

The current research proposes that online ratings happen similarly to the experiments performed by Tversky & Kahneman (1981). That is, online users are presented with information they can assess (i.e. a decision problem). However, different SNSs and CCc show this information differently. That is, different social media platforms frame UGC through the type(s) of media they allow, the number of characters of each post, the rating system available, requiring user profiles or not, etcetera. Hence, users presented with the same information on different websites might rate it differently depending on how the content is displayed on that particular site. Therefore, it is proposed that the choice-making process of surveys depicted in Figure 2.7 (Tourangeau et al. 2000) – which has been adopted by scholars researching online ratings (e.g. Riedl et al. 2010; Riedl et al. 2013) – should include the concepts of frame and biases (Kahneman & Tversky 1979; Tversky & Kahneman 1981; Kahneman 2003). Figure 2.8 presents this proposal.

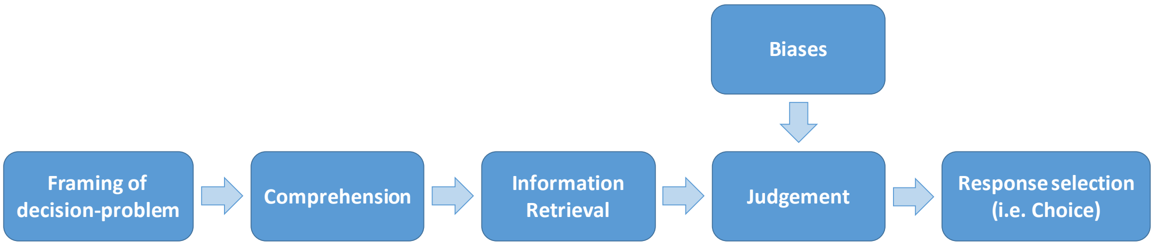


Figure 2.8 - Proposed choice-making process in social media

2.3.3 Heuristics and biases

Following their experiments, Tversky and Kahneman outlined three heuristics: representativeness (i.e. the way people infer probabilities from different samples), availability (i.e. the ease or effort at which a particular concept comes to mind), and adjustment and anchoring (i.e. when the estimate of a starting value is given by a previous number), (Kahneman et al. 1982). Those heuristics gave place to the study of other systematic errors (i.e. biases) that take place when people make choices. The study of biases has fascinated academics across a broad range of disciplines, and there are over a hundred listed biases supported by research (MacLachlan 2014).

It could be said that biases are somehow required for the transmission of knowledge. From an individual perspective, people can only process a limited amount of information because of the need to make fast, simple, and unconscious decisions that have enabled the species to survive (Kahneman 2012). Moreover, from a social perspective, the human species has evolved to allow individuals to forego the costs of individual learning, enabling humans to acquire knowledge through social learning processes like teaching, language and imitation (Mesoudi 2011). Combining both perspectives, it could be claimed that the human species has evolved towards implementing mental structures that frame – and possibly bias – information, so as to quicken, simplify, and reduce the number of choices people must make.

Moreover, if individuals relied on social processes that simplified the transmission of information even when people were in relatively small groups, that need is even greater now. With almost half of the world's population online (Internet Live Stats 2018), and "as many digital bits as there are stars in the universe" (Turner et al. 2014, p.1), website designers have had to find a way of creating heuristics that would allow users to browse content and make choices as efficiently as possible. As Park & Nicolau (2015) have stated, consumers are seeking heuristic information cues to simplify the amount of information involved in taking decisions within online environments. For instance, as of today, there are 18,373 restaurants listed in London on the travel site TripAdvisor (2018). This means that if a user wanted to browse the individual websites of each restaurant, they would need to invest 919 hours (38 days without any break), assuming they were to spend just three minutes collecting relevant information regarding the type of cuisine, location, prices, and so forth. In the end, even if someone dedicated all this time to looking at restaurants, they would probably not be able to make a choice: firstly because most websites would claim to have the best food, location and prices; secondly because the sheer number of possibilities would overwhelm and paralyse them (Schwartz 2009). Thus, it can be argued that users need sites where they can find simple and reliable information to make quick choices. However, this can only be accomplished through website designs that, by simplifying the sharing of UGC, also frame the content and make it prone to biased transmission.

As previously mentioned, there are more than a hundred biases supported by research, and most could potentially be applied (or have been applied) to research done in online environments. For instance, there have been studies regarding the difference in choices when two alternatives are presented simultaneously or sequentially (Ariely et al. 2006), others regarding the timing of decision-making (Ariely & Zakay 2001), and still others concerning the emotional state of the decision-maker (Andrade & Ariely 2009). However, the focus of the present research is to understand choice-making through identity, groups and relationships; and these concepts need to be studied in a social context. Moreover, the reader is reminded that both the definitions of social media and knowledge used in this thesis highlight the importance of the community aspects of both terms. Therefore, the biases that have been used are also social in nature; that is, they are biases that potentially take place when individuals are in groups.

2.3.4 Group biases: content, prestige & conformity

The group-biases that were chosen for this research have been used by several scholars in the field of cultural evolution. Their main argument, similar to that of Simon (1955; 1979) and Tversky and Kahneman (1974; 1981) is that individuals do not acquire knowledge by assessing the payoffs of every piece of information they come across. Instead, people rely on social learning or knowledge transmission to acquire most of their beliefs (Durham 1991; Henrich 2001; Richerson & Boyd 2005; Mesoudi 2011). Nevertheless, people do not merely acquire random beliefs from random people. Instead, the social learning process in which knowledge is shared contains a number of biases; which is why it is known as 'biased knowledge transmission' or 'knowledge selection'

(Durham 1991; Henrich 2001)[10]. The group biases that relate to choice-making fall into three categories. The first class, *content biases*, refers to the selection of a belief because of specific qualities of the presented information that make it more likely to be chosen. Conversely, the other two groups have to do with the context of the information, rather than with the characteristics of the belief itself. The second group, *prestige biases*, result when people acquire beliefs from individuals with specific characteristics (e.g. successful, prestigious, or similar to themselves). Finally, the third category, *conformist biases*, happen when someone imitates the beliefs that are expressed by the majority of the group (Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011).

It should be noted that scholars who have used quantitative modelling techniques, such as Cavalli-Sforza & Feldman (1981) and Boyd & Richerson (1985), have received criticism from researchers in other disciplines that study culture and knowledge evolution from a qualitative viewpoint. Two of the most common critiques arise from models being based on assumptions that deal only with the individual-level, in addition to being too similar to biological evolution (Durham 1991; Mesoudi 2011). Nonetheless, the same scholars who have criticised these models also acknowledge that they do have the advantage of simplifying processes and tracking them over time under many conditions, which could not be done otherwise. Thus, quoting Mesoudi (2011, p.57), "quantitative models generate clear predictions that can be tested in the lab or the real world". More on these issues will be discussed in the methodology chapter. However, it is important to outline that these group-biases have not been studied in the way of a model here, nor has this research assumed a neo-Darwinian position, as outlined at the start of this chapter. Instead, this study is based on how researchers have previously applied, observed and studied these biases empirically in real-life situations, within a social context (e.g. Mesoudi 2011; Richerson & Boyd 2005). In the same manner, the researcher has attempted to observe the biases in online settings, in the context of groups and relationships.

In the following pages, each category of biases is explained in-depth, together with how they are envisioned to take place on social media sites. Chapter 3 further discusses these

[10] Note that the cited scholars (i.e. Durham 1991; Henrich 2001; Richerson & Boyd 2005; Mesoudi 2011) use the term 'culture' instead of 'knowledge'. As mentioned, culture is seen as being too broad, whereas knowledge is more specific.

biases in the context of the experiment, and Chapters 4 and 5 describe precisely how they are measured. Thus, the goal of the present chapter is to show and compare the different ways in which the three group-biases might occur in SNSs and CCs due to their different designs. Moreover, it is of relevance to show examples of the most used sites at this point, as these provided the initial inspiration for the design of the study. However, as previously explained, the experiment takes place in *one* platform, which was modified to imitate the different designs of SNSs and CCs in terms of user profiles and rating scales.

1) Content-based biases

This group of biases, also known as 'direct biases', result from the interaction of particular features of a belief with people's social learning. That is, the presence of certain features make individuals more (or less) likely to select the content (Henrich 2001). In the broader world of culture, an example commonly given is food. Individuals evolved into surviving through increased consumption of fatty foods, and thus people are more likely to choose meals that contain fat because of this (Henrich 2001). Yet, in the field of beliefs, examples become more complex given that individuals' cognitive structures are responsible for making some content more easy to be learnt, remembered or acquired (Richerson & Boyd 2005). For instance, Mesoudi (2011) exemplifies these biases with the diffusion of rumours, where it has been found that those provoking strong emotional reactions of disgust are more likely to be selected. Interestingly, this assertion was tested in online environments through the selection of memes, where researchers studied choice and transmission of urban legends of the internet through the content's propensity to evoke emotions such as disgust, fear and anger (Heath et al. 2001). The results of the internet experiment showed similar results as those from the dispersion of rumours, concluding that people would be more likely to make choices based on 'emotional selection'.

Drawing from the description of content-based biases outlined by Henrich (2001), Richerson and Boyd (2005) and Mesoudi (2011), content biases occur in social media when users select UGC (e.g. a post, review, video, image) because of particular aspects of the content that are considered advantageous. Of course, this would depend on the type of media that is allowed on each site, and its aim. For instance, it is possible that selection due to high emotions takes place in SNSs, where people share day-to-day stories of their lives. For example, if a study was conducted on Facebook, content biases could be measured by the sentiment of posts. Nevertheless, this again would depend on the type of social media site, as some elements might be relevant for some and not for others. For instance, the presence of a hashtag (i.e. a word or phrase preceded by the hash symbol, which serves to differentiate topics on social media) has been shown to be significant for the selection of UGC on Instagram and Twitter, whereas it is not the case for Facebook. Regarding Instagram, hashtags related to food are attractive to users (Hu et al. 2014); whereas in Twitter it has been found that social contagion happens at higher rates when tweets are accompanied by hashtags containing uncommon words or are politically controversial (Cunha et al. 2011; Romero et al. 2011).

Furthermore, it may not be relevant to investigate if the length of tweets is a content bias, because these are more or less the same length, as they cannot surpass the 140 characters. In contrast, the length of reviews has proved to be a significant aspect in the selection of travel advice, among other elements. For instance, scholars researching travel CCs have found that the length of reviews and their readability can predict if other users will find their content useful or not (Liu & Park 2015; Cheng & Ho 2015; Park & Nicolau 2015). In addition, extreme ratings are perceived as being more useful, independently of them being positive or negative (Park & Nicolau 2015; Jalilvand et al. 2011). Therefore, according to the given ratings of users (i.e. useful votes, helpful votes, likes, or thumbs up), the most up-voted reviews are long, easy-to-read and are not neutral. Figure 2.9 shows an example of aspects that could be considered content biases in TripAdvisor. As can be seen, according to the previously outlined research, specific tags utilised by users, extreme ratings, or the length and readability of each review could be used as predictors for other users acquiring this knowledge, which would be inferred by their use of the helpful/thumbs-up rating.

Nevertheless, it should also be mentioned that academics studying online reviews have found that the factors that have the most impact in determining the usefulness of reviews have lesser to do with content-based elements. Instead, ratings show a higher association with the personal characteristics of the reviewers, such as the presence of their real name or profile picture, expertise, number of ‘fans’ or ‘followers’ (Liu & Park 2015; Park & Nicolau 2015; Cheung & Thadani 2012; Cheng & Ho 2015). These findings from social media research suggest that content-based biases are weaker in comparison with those related to the characteristics of the individuals authoring the information (i.e. prestige-based biases).

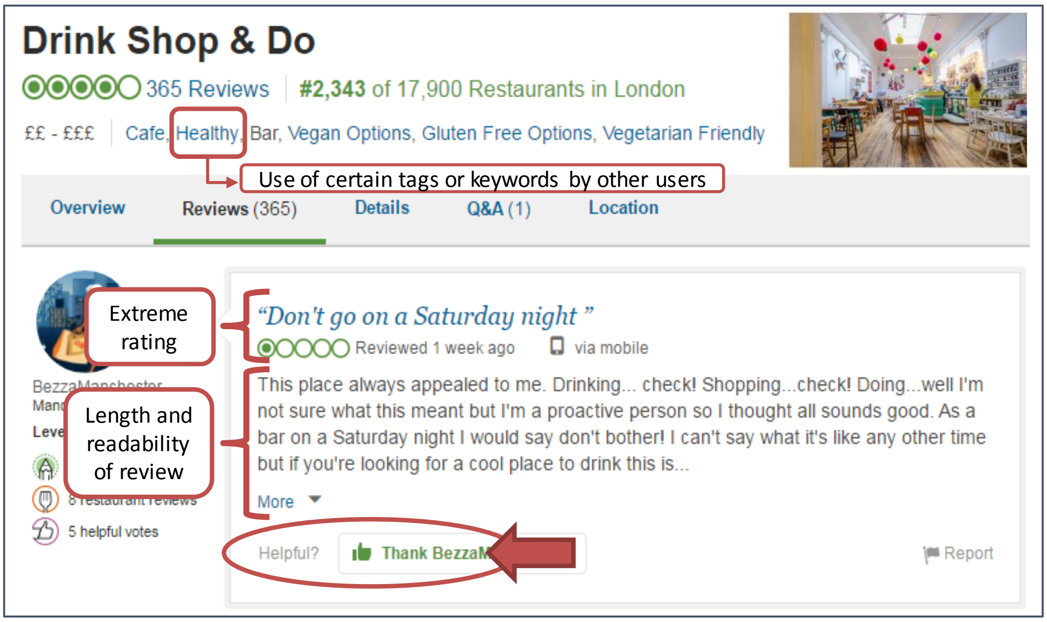


Figure 2.9 – Example of content-based biases on TripAdvisor

2) Prestige-based biases

This class, also known as 'model-based biases', take place when individuals imitate the observable qualities of the individual who is sharing the content. This happens because people have a predisposition to imitate successful people and those similar to themselves (Richerson & Boyd 2005). Prestige has been linked to different personal characteristics. For instance, some researchers have associated it with status, skill, knowledgeability, and age; that is, older individuals or those with higher social status or more skills would be considered more prestigious and therefore more people from their communities would copy their beliefs (Henrich & Gil-White 2001). Moreover, these biases can occur even when the content has nothing to do with the area of expertise of the person who is transmitting it. An example was Michael Jordan's recommendation of a fragrance, where individuals followed his advice even though his excellence in playing basketball was not affected by the scent he used (Henrich & Gil-White 2001; Henrich 2004).

Further evidence from prestige biases arises from various offline studies in social psychology, where participants consistently align and even change their choices to imitate individuals who are deceitfully introduced as having a certain level of power or expertise (e.g. expert professor, art director, successful horse bettor, etcetera) (Mesoudi 2011). In addition, organisational theorist Karl Weick (1969) stated that the best thing for a network was to follow individuals with the best practices. Indeed some real-life experiments on knowledge and innovation diffusion have also confirmed the presence of prestige biases through what is known as ‘opinion leaders’, who are described as influential individuals within a network (Cross, Parker, et al. 2001; Iyengar et al. 2011). Likewise, online research on opinion leadership has further supported the presence of prestige biases (Litvin et al. 2008; Wu et al. 2011; Jacobsen 2015). Specifically, recent studies on electronic word-of-mouth (eWOM) found that users were more likely to consider useful the reviews of individuals who have more followers, and higher levels of expertise, ‘reputation’, and valuable or helpful votes (Liu & Park 2015; Cheng & Ho 2015).

The previous paragraph highlights some elements that differentiate prestige in online and offline environments. As mentioned in the first section of this chapter, in real-life, users are subject to their physical presence. However, the internet and computer-mediated communication (CMC) have made disembodiment possible (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008). Thus, in offline environments, an assumption would be that people know specific characteristics of the members of their groups, such as age, gender, and whether they are skilful, knowledgeable or successful. However, in most cases, it is not possible to know these characteristics in online environments. As a result, social media sites have adopted designs in which, through badges and votes (i.e. *gamification*), they give ‘online prestige’ to users, at the same time that they try to increase their engagement (Denny 2013; Robson et al. 2016). Therefore, it could be argued that prestige-based biases in SNSs happen similarly to real-life, as users know each other offline, in most cases. Conversely, as anonymity is more common in CCs, users would be motivated to obtain ‘online prestige’ through votes and badges, to make this information part of their online identity.

Hence, the way in which this group of biases was studied in this thesis is by differentiating between real-life and online-gained prestige. The former comes through individuals being successful, skilful or having characteristics that would make others more prone to acquire beliefs from them; the latter involves gaining votes or badges within a social media site.

a) *<u>Real-life prestige: Exalting successful individuals</u>*

People who are already successful in real-life are sometimes given special profiles within SNSs and CCs. In this way, other online users can be assured that these accounts are genuinely those of the individuals whom they have heard about offline. Figure 2.10 for example, shows the ‘verified’ profiles of the ex-president of the USA, Barack Obama. Users can recognise that it is the real profile by the blue ticks that both Facebook and Twitter provide. Based on the research mentioned above, it would be safe to assume that users would be more prone to acquiring UGC from these individuals by virtue of the status accorded them by the site.

Figure 2.10 – Example of real-life prestige: verified profiles on a SNS and a CC

However, it should be highlighted that ‘normal’ individuals can acquire prestige for different reasons. For instance, it has been found that, on certain social media sites, people with specific nationalities receive more attention than others. Namely, a study on Twitter found that Indian users were aware of what was trending in the USA, whereas the converse was not true (Wilkinson & Thelwall 2012). Finally, research performed in online reviews within travel sites showed that reviewers who display their real names, profile pictures, and addresses, received more useful votes (Park & Nicolau 2015; Liu & Park 2015). These studies did not catalogue declared or perceived nationalities, but it can be argued that by increasing self-presentation, users differentiated themselves from the millions of users who were unknown, and this gave them more prestige.

b) Online-gained prestige: Creating opinion leaders through votes and gamification

As previously outlined, online-gained prestige is usually obtained by acquiring status through up-votes, badges and followers. This, of course, can reflect real-life characteristics such as a high level of expertise on a particular topic or a higher engagement than other users within the network. In SNSs, online-prestige is usually achieved by having a high number of followers. Conversely, most CCs have a mechanism to vote for other users and earn badges, and therefore rank them depending on their level of expertise. For instance, Figure 2.11 shows the summary of a users' online-gained prestige in TripAdvisor. This CC ranks its users by level of contribution[11] from one to six, depending on the number of reviews (total and by category) and by the sum of received helpful votes.

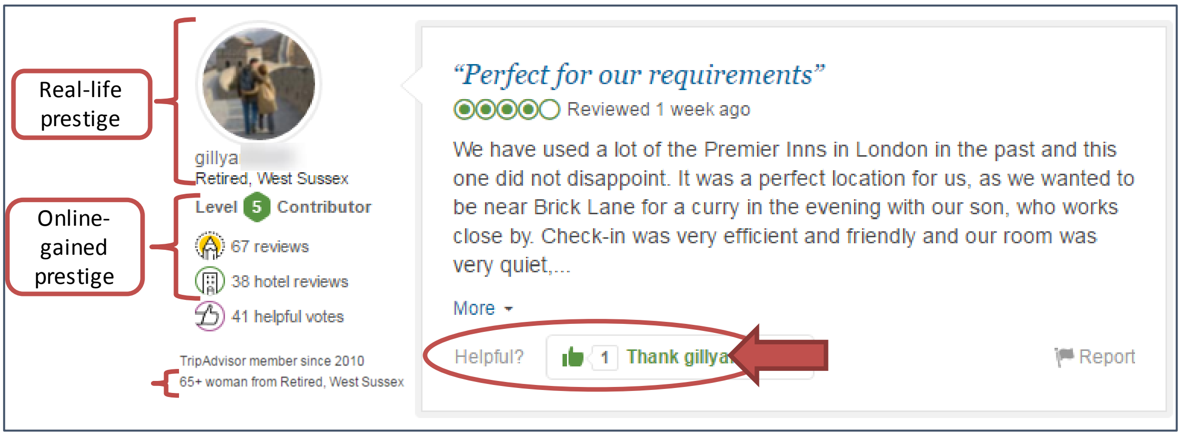


Figure 2.11 – Example of prestige-based biases on TripAdvisor

To sum up, with the example shown above, if someone acquired a belief (i.e. gave a helpful vote) from someone because of their level of contributor or number of reviews, it could be said that it was due to online-gained prestige. Instead, if this were because of her declared age, gender or picture, it would be considered real-life prestige. Moreover, in both cases, it would be catalogued as a prestige-based bias. As has been mentioned, each website has its unique design so sites like Stack-Overflow, Amazon, Google, among others, have other online-prestige mechanisms in place.

3) Conformity-based biases

This category, also known as frequency-based biases, refers to the commonness or uncommonness of the different content from which people can choose (Richerson & Boyd 2005). Information that is chosen due to these biases is not assessed regarding its advantages or disadvantages, but rather due to conformity (or anti-conformity) with the rest of the population (Mesoudi 2011). Hence, these biases occur when individuals prefer

[11] This ranking has evolved during the course of this PhD. Before the summer of 2015, TripAdvisor would rank its users in the following way: reviewer, senior reviewer, contributor, senior contributor, or top contributor.

to imitate the beliefs that are followed by the majority and not the minority of the group, and this holds true even when other group members do not know the choice made by the person acquiring the belief (Henrich 2001). Regarding evolutionary models, these have shown that conformity biases are very powerful in transmitting knowledge (Boyd & Richerson 1985; Henrich & Boyd 1998; Henrich 2001; Henrich 2004; Richerson & Boyd 2005; Perreault et al. 2012).

Likewise, lab and real-life experiments have shown similar results. For instance, social psychologists have conducted experiments in which they placed an individual within a group of people whom had received specific instructions from the researchers. Everyone was asked to make a very simple and obvious choice and share it with others; in most cases, participants wrongfully aligned their choices to the majority of the group (Asch 1956). Years later, the experiment was repeated – increasing the difficulty of tasks and including incentives for accuracy – finding that this increased conformity and social influence within groups (Baron et al. 1996). Similarly, conformity has been used as a plausible explanation as to why some groups become homogeneous while – at the same time – creating higher between-group differences; which has helped explain why, when migration takes place, diverse groups adhere to their own traditions even when sharing similar weather and landscapes (Richerson & Boyd 2005; Mesoudi 2011).

However, despite these studies, the following issue has been raised: "Conformity does not stir much interest among contemporary social psychologists; the work conducted between 1950 and 1980 is still the main stuff on modern textbooks. Conformist *transmission* remains very poorly studied […] Without Darwinian concepts and tools, the population-level consequences of individual behaviour are not intuitive […] Understanding rather precisely how *individuals* deploy their kit of imitation heuristics is necessary to understand the rates and direction of [knowledge][12] evolution, and work on the problem has hardly begun." (Richerson & Boyd 2005, pp.123–4). Although this quote is more than a decade old, it can be argued that it stills holds true, especially within online environments, as will be explained in the following paragraphs. Therefore, to address this gap (LG-6), the present thesis has included the study of conformist transmission with a 'Darwinian perspective', i.e. through the analysis of the VSR mechanisms.

[12] The original phrase reads 'culture' evolution

Furthermore, the way in which conformity-based biases were thought of taking place in social media is as follows. Almost all sites display (or frame) UGC in a way that shows how many users have previously selected it (i.e. liked, disliked, up-voted, down-voted, or simply rated it in a particular way). For instance, using the same example of TripAdvisor (seen in Figure 2.9, explaining content-based biases; and Figure 2.11, regarding prestige-based biases), Figure 2.12 now shows how conformity-based biases are thought of taking place in this CC. As seen from the image below, if a user were more prone to acquiring a belief because of other users having done so before, this would be due to conformity.

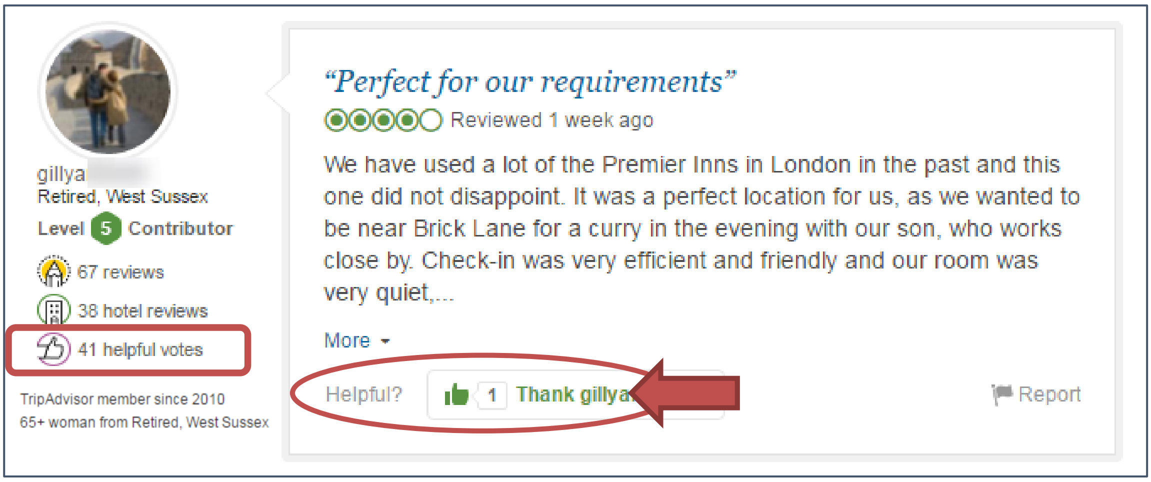


Figure 2.12 - Example of conformity-based biases in TripAdvisor

However, most studies performed in SNSs (Wu et al. 2011; Wilkinson & Thelwall 2012) or CCs (Liu & Park 2015; Park & Nicolau 2015) fail to analyse conformity, probably because it would be complicated and more time-consuming to determine the effect of the group on the choices of individuals. Likewise, controlled experiments studying users' decisions through ratings in 'artificial' online environments have designed experiments in a way that ratings are independent of each other, trying to avoid "confounding effects through biasing cross-influences between rating scale users" (Riedl et al. 2013, p.9; Riedl et al. 2010). Regarding this last point, the author would argue that there is very little difference between an offline and an online experiment if the ratings provided by other users are not displayed, given that this is a fundamental characteristic of online sites. Thus, there is a lack of research addressing conformity in online environments, which seems to be a combination of theoretical, empirical and methodological gaps (LG-7). Therefore, the following section can be seen as a proposal and rationale for studying conformist transmission in social media sites.

2.4 STUDYING CONFORMITY IN SOCIAL MEDIA THROUGH GROUPS AND RELATIONSHIPS

This thesis has embraced the principle that identity is not an individual product but rather a social construct; consequently, self-presentation and self-disclosure adapt to different social contexts (Goffman 1959; Goffman 1963; Suler 2002; Schau & Gilly 2003; Suler 2004; Meng & Agarwal 2007; Zhao et al. 2008). Likewise, social media was defined as a communication medium for individuals within social structures (Peters et al. 2013). Therefore, given that this research is studying the transmission of information within networks, it makes sense to ask one of the maxims of communication: “who says what, in which channel, to whom, with what effect” (Lasswell 1948, p.117). To answer this question and understand the way in which UGC is transmitted within networks, the research also needs to take into account the effect that groups and relationships have on communication.

As argued in the preceding section, there is a need for more studies to deal with conformity in online environments. However, there is arguably also a need for a different way to study conformity in social media. Previous studies have assumed that all members of a group know each other (e.g. Henrich & Boyd 1998; Henrich 2004; Richerson & Boyd 2005; Perreault et al. 2012). However, although this may be true for some communities, it is far from the reality of online environments. With almost half of the world’s population online, it would not be even remotely possible to suppose that all users know one another. Therefore, this section firstly advances the need for a distinction between the whole and personal networks. After that, it suggests a further differentiation within the personal network, based on the strength of relationships. Finally, it summarises how the study of conformity should take place in online environments.

2.4.1 Groups: differentiating between the whole and personal networks

In the current era of fake news, Facebook and Twitter bots, exaggerated marketing campaigns where ‘likes’ and ‘followers’ can easily be bought, it is understandable that online users exhibit lower levels of trust towards politicians, marketers and unknown users. Indeed, research has found that millennials – young adults born between 1980 and 1995 and the most avid users of the internet – trust the UGC from their friends more than information (e.g. posts, reviews, sales ads) that comes directly from companies (The McCarthy Group 2014; Gutiérrez-Rubí 2014). This may be a reason why more sites are being re-designed to make it easier for users to sign-in with their Facebook or Google+ profiles, where most people are identifiable at least to some of their offline friends.

On the one hand, differentiating among those who are known to them in real-life might be beneficial for online users given that people evolved by foregoing the costs of individual learning and by relying on others through teaching and imitation (Mesoudi 2011). Therefore, by requiring people to sign-in with their real identities, websites allow users to quickly detect those whom they already trust. On the other hand, this also enables websites and companies to 'exploit' peoples' personal networks by arguably converting everyone into a marketer. A few examples of sites that have integrated with SNSs are Quora, TripAdvisor, Netflix, Spotify, Instagram, and Amazon, among others. Thus, sites that allowed anonymity are now strongly encouraging users to disclose to their networks where they are dining, what they are asking, watching, listening to, and buying. In the same manner, people nowadays can know all these facts about their friends and acquaintances. The constant presence of others might have an impact on people's identity and, consequently, in the way they perceive information and make choices.

Regarding individual's identity, it has already been explained that people would present different 'selves' depending on the audience and context (Goffman 1959; Goffman 1963). Therefore, by encouraging that users sing-in with their SNSs' accounts, CCs are increasing users' self-presentation. Furthermore, just as identity is context-dependent, a number of social psychologists claim that identities are subject to groups (i.e. *social identity theory*), attesting that these vary in a continuum between acting in terms of the self (i.e. interpersonal) or of their group (i.e. intergroup) (Tajfel 1974; Tajfel 1978; Tajfel & Turner 1979). Also, they argue that it takes very little for people to cluster into groups, in which case they tend to favour the *ingroup* as opposed to the *outgroup* (Tajfel 1974). This happens because people struggle to maintain a positive self-concept and, as this is linked to their association with a determined social group, they strive for a positive social identity which is achieved by favourable comparisons between the ingroup against relevant outgroups; a concept known as *ingroup bias* (Tajfel & Turner 1979). These concepts later evolved into what is referred to as *self-categorisation theory*, which builds from the idea that people sometimes define themselves as individual or group entities.

When in a group, people try to self-categorise by minimising the existing differences between them and other ingroup members, while separating themselves from those in the outgroup (Turner 1984; Turner & Reynolds 2012).

Hence, if individuals within groups adapt their identities, it can be expected that the way they perceive information, and consequently their choices, are also affected. Moreover, as both social identity and self-categorisation theories outline, if people within groups decrease the perception of the differences they have with other ingroup members, it can be expected that their choices will become more similar. Therefore, referring back to the examples of social media shown above (Figure 2.9 to Figure 2.12), perhaps being able to differentiate their real-life ingroups online makes users more prone to access the information of those whom they already trust and see as being like-minded. As a result, the UGC they access from their ingroups might also be perceived as better suited to them, as the apparent differentiation is diminished. Thus, they might end up choosing similar music, restaurants and movies to those chosen by other members of their ingroup.

It should be highlighted that there are both theoretical and empirical gaps to be addressed in relation to the presence of groups in online environments (LG-8). First, from a theoretical perspective, it has been pointed out that there is a need to explain the communication shifts that take place when offline groups are transformed to online groups (Kietzmann et al. 2012). Moreover, as noted above, most scholars studying online reviews and ratings ignore the effect of conformity towards the overall network (e.g. Riedl et al. 2010; Wu et al. 2011; Wilkinson & Thelwall 2012; Riedl et al. 2013; Liu & Park 2015; Park & Nicolau 2015). That said, it should be mentioned that there is one recent paper that deals with the effect of 'social approval cues' on decision-making (Mueller et al. 2018). In this research, approval cues towards a business idea were formulated through the number of contributors, the percentage of requested funding awarded, and the sum of Facebook 'likes'. Two ideas were presented to participants: one with high and one with low social cues. However, a valid question would be if participants would have taken different decisions when presented with ideas with low social cues if the Facebook likes displayed had all been by people from their ingroup, or if they had been told that the author of the presented idea was someone whom they knew. For all these reasons, and to fill in the outlined gaps, the present research differentiates between the whole network (i.e. outgroup) and personal networks (i.e. ingroup).

2.4.2 Relationships: taking personal networks a step further through the strength of ties

It has been alleged that a system depends not only on the elements that it comprises but also on the direct and indirect connections among those elements (Loasby 2002). However, not all their connections are necessarily equal. For instance, the frequency and depth of conversations between oneself and a close friend or an acquaintance may vary significantly. The strength of relationships has long been studied in the field of *social networks* and has been applied to a range of fields including consumption of products, the spread of diseases, politics, criminology, and job seeking, among many others (Jackson 2008). Most importantly – and what interests the present study – is that social networks and the strength of relationships play a crucial role in the transmission of information (Jackson 2008). Thus, deepening the analysis by studying relationships within groups can improve understanding of Lasswell's (1948) maxim cited above, by addressing 'who says what to whom, with what effect'. Answering this question has been of great interest for scholars studying 'offline' social networks (e.g. Cross, Parker, et al. 2001; Cross, Borgatti, et al. 2001; Cross et al. 2002; Dobson et al. 2013), as well as online environments (e.g. Wu et al. 2011; Wilkinson & Thelwall 2012; Jiang et al. 2014).

It is essential to highlight that the differentiation of relationships is something relatively new in social media. That is, although it was always possible for users to identify those known to them offline as long as they were non-anonymous, sites did not have a design that allowed for this distinction to be made 'official'. Therefore less than a decade ago, scholars alleged that "social media treats all users the same: trusted friend or total stranger, with little or nothing in between" (Gilbert & Karahalios 2009, p.211). However, this statement would be considered false in the present day. In mid-2011 Google introduced its SNS, Google+, which presented an innovative idea called 'circles' where users could create groups based on the strength/type of their relationships and share targeted content with each group (SEO Web Marketing 2012). By the end of the same year, Facebook introduced 'smart lists' to manage friends and, similar to Google+, allowed users to share and view specific content with each list (Loomer 2012). The only difference is that Facebook created and updated specific lists for users, based on location or declared job or family relationships (Vahl 2011). Since then, the presence of differentiated personal networks has appeared not only in these two SNSs but also in the many CCs where users are strongly encouraged to login using their Facebook or Google+

accounts. Concrete examples are presented in the next subsection but, for now, it is relevant to mention some ways in which the strength of relationships might have an impact on identity and groups, and consequently on the transmission of information.

Concerning identity, as has already been mentioned, people present themselves according to their audience (Goffman 1959; 1963). Therefore, it becomes evident that need for ‘circles’ and ‘smart friends’ lists’ is at least partly to avoid sharing certain information with the wrong audience and therefore incur in ‘embarrassments’ (Goffman 1959; Kietzmann et al. 2012). Likewise, having explained that individuals adapt their identities based on their group memberships and link their self-esteem to these (Tajfel & Turner 1979), it is understandable why people would like to differentiate among distinct groups within social media. Therefore, differentiating among different relationships seems like a natural step in SNSs because these sites rely on social ties and people naturally want to share and receive more information with those with whom they have stronger friendships, as happens in offline environments. However, CCs revolve around content so, when they highlight to people who their close friends are, this might make those individuals more prone to only communicating with others who they already know and are most similar to them, therefore creating the so-called ‘echo-chambers’. Hence, it could be argued that, as the ingroup-outgroup presence generates ingroup biases (Tajfel & Turner 1979), people’s choices might also get affected in this process.

The rationale for the previous argument is reinforced by *social network theory*. In relation to the sharing of information, scholars in the field of social networks have found that “a significant component of a person’s information environment consists of the relationships he or she can tap for various informational needs” (Cross, Parker, et al. 2001, p.100). The different types of relationships have been found to affect the transmission of information in a number of ways. However, before describing this issue, two concepts should be defined. First, *networks* can be defined as “a set of actors connected by a set of ties” (Borgatti & Foster 2003, p.992). Further, the *strength of a tie* has been defined as “a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterise the tie” (Granovetter 1973, p.1361). Thus, Granovetter (1973) initially classified ties as *strong* (i.e. relationships that involve larger time commitments and have a significant overlap in their friendship circles); *weak* (i.e. friendships characterized with less similarity and contact-time, which have an intermediate overlap in their friends' circles); and *absent* (i.e. where there is a lack of relationship between two actors).

There are two properties to highlight from these terms. First, it has been found that the transmission of complex knowledge, sharing, and the strength of collective beliefs are favoured by stronger and closer connections, that is, by stronger ties (Dobson et al. 2013). Second, stronger ties require larger time commitments, and therefore two close friends of a determined person are likely to know each other and hence have a strong relationship as well (Granovetter 1973). That is, strong ties lead to groups of very similar people, a concept denoted as *homophily* (Kilduff & Brass 2010; Peters et al. 2013). This creates a paradox in that people with a high number of strong ties but few weak ones "will be deprived of information from distant parts of the social system and will be confined to the provincial news and views of their close friends" (Granovetter 1983, p.202). Thus to recap, the importance of strong ties relies on the transmission of complex knowledge and collective beliefs, whereas weak ties are essential for the diffusion of information.

Given that both strong and weak ties play an important role, an important question arises. There seems to be a common belief that online environments make possible the communication of virtually anyone on the planet, but could this start changing with individuals being shown primarily the information of their close friendship groups? As previously outlined, SNSs only began to allow the segregation of friendship groups by strength towards the end of 2011. In addition, the integration between SNSs and CCs began in mid-2013. Therefore, at the time of writing this thesis, users have experienced websites highlighting the content of their closest friends for less than five years, which means both users and researchers have had very little time to measure the effects of this phenomena. Also, with both SNSs and CCs 'closing the doors' to their APIs[13], it has become very difficult for academics to measure the effects that people's personal networks and the strength of ties are having on the transmission of UGC and on peoples' choices. Consequently, there are a few gaps to be addressed. For instance, it has been listed as a theoretical gap (LG-9) that there is a need for joining a social media model with real-life data (Gilbert & Karahalios 2009). Moreover, linking the three building blocks – relationships, groups and identities – highlights the need to answer the following:

[13] Application Programming Interface (i.e. the 'communication' or 'access' to a website).

how do online ties relate to those offline? Specifically, could ‘online embarrassments’ have negative results in offline relationships? (LG-10), (Kietzmann et al. 2012).

2.4.3 Including groups and relationships to the study of conformity in social networking sites and content communities

Having outlined the need for studies that differentiate between conformity to the whole and to personal networks, and the strength of ties within the latter, this section outlines how these could be studied within online environments. It should be noted that, just as in the previous section on biases, the following examples are from real SNSs and CCs. Nevertheless, for scholars to perform research on these sites, they would need access to each of the participants’ online ingroups and the classification of their friendships. Conversely, if they wanted to imitate these properties in other sites, they would need to know the subjects’ real-life friendships. This increases the complexity of the research, and although it can be difficult, there is the need to study if the way people access information and make choices are affected by the different ways in which SNSs and CCs use people’s personal networks. Chapters 3 and 6 explain in detail how groups and relationships are studied in this thesis; for the time being, it is worth examining and comparing how different sites manage users’ personal networks.

Conformity to the whole network

As previously mentioned, conformity biases occur in online environments when a user selects a particular piece of UGC (e.g. post, review, video) because other users have previously done it. Moreover, when conformity is to the ‘whole network’ (e.g. Granovetter 1983) it means that an individual imitates people from the SNS or CC, regardless of the relationship (s)he has with them. This could happen when websites do not differentiate among different strengths of ties or when users only consider the frequency of other individuals within the SNS or CC that ‘agree’ with a determined belief, regardless of whether they know them in real-life or not. Arguably, SNSs and CCs make users prone to acquiring information based on conformity to the whole network by highlighting the total number of people who have already selected it. Figure 2.13 shows how Facebook highlights the members of the network that have ‘liked’ specific content.

Figure 2.13 – Example of conformity to the whole network, Facebook

Likewise, CCs also highlight how members of the network have selected information. However, as most of them allow at least a binary rating scale, users are aware of conformity as well as anti-conformity within the whole network. Figure 2.14 shows an example of *YouTube*, which has a dichotomous rating scale, while Figure 2.15 presents one from *TripAdvisor* which uses a 1 to 5 likert scale. It should be noted that the three displayed figures are representative of the most commonly used rating scales in SNSs and CCs: one-point, dichotomous, and likert. Moreover, as outlined in the first section, they all provide for different levels of self-disclosure. Thus, it should not be forgotten that self-presentation, self-disclosure, groups and relationships all play a role in the choice-making process of users.

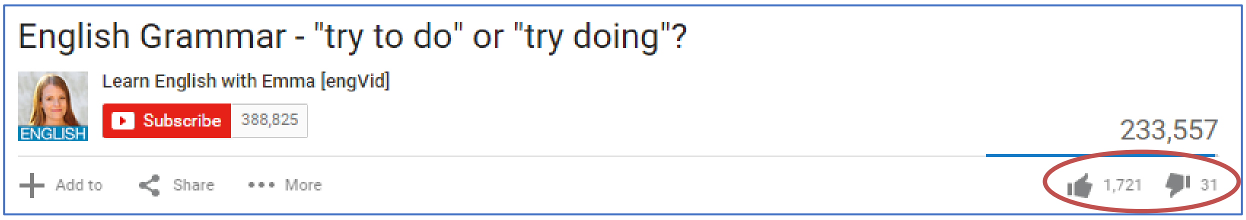


Figure 2.14 – Example of conformity to whole network with dichotomous scale, YouTube

Figure 2.15 – Example of conformity to whole network with likert scale, TripAdvisor

<u>*Conformity to personal networks (i.e. ingroup = present ties)*</u>

Conformity to personal networks happens when users choose information previously selected by people they know in real-life. This type of conformity occurs when SNSs or CCs show users that a post, image, or video has been liked, disliked, or rated in a particular way by someone within their personal networks (it should be noted that the personal network is not yet differentiating among different tie strengths).

As mentioned, the presence of personal networks happens more commonly on SNSs as they rely heavily on social bonds. The best example is Facebook, where most people's 'Facebook friends' are also friends in real-life (Quan-Haase & Young 2010). When UGC is shared in this site, the website not only displays how many people have 'liked' it, but also highlights how many of these belong to the user's personal network, as seen in Figure 2.16.

Figure 2.16 – Example of conformity to personal network, Facebook

Conversely, CCs do not rely as heavily on personal bonds because the focus is on content and members are not assumed to know each other outside the website. However, as discussed above, in recent years communities like YouTube, TripAdvisor, Netflix, Quora, Coursera and many others have started allowing – and encouraging – users to sign-in using their Facebook and Google+ accounts. This, the researcher would argue, has made the line that divides SNSs and CCs very thin, as it increases the levels of self-presentation and therefore makes users in CCs more likely to rely on personal networks. At the same time, online users are now constantly aware of the preferences of their real-life friends regarding videos, restaurants, movies, online courses, and many more. For instance, Spotify users who have signed in with their SNS account can see what their friends are listening to, and they can select this content (i.e. music) with just one click, as seen in Figure 2.17.

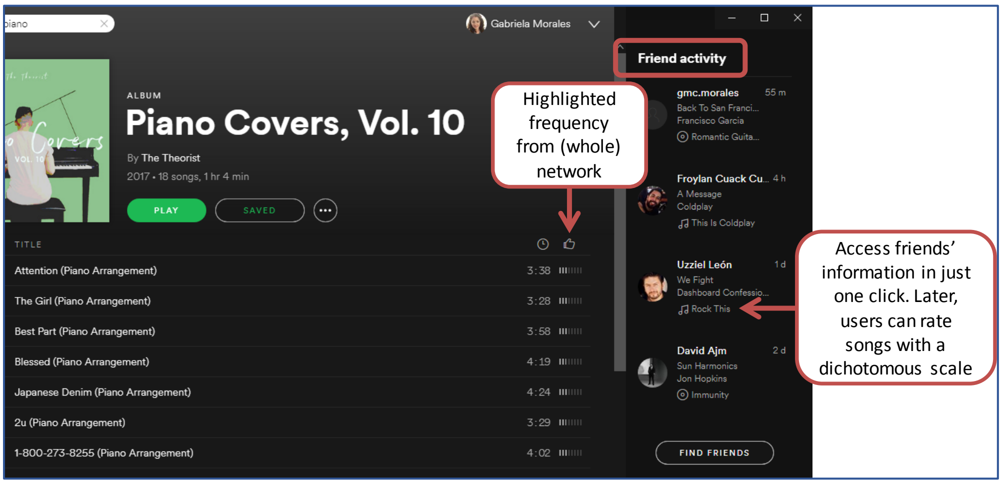


Figure 2.17 - Presence of personal network in Spotify

Likewise, TripAdvisor highlights to users where their friends have been, allowing access to their reviews (see Figure 2.18). Hence, in the first example users can see the information that their personal networks are 'consuming', whereas in the second one, the content (i.e. reviews) are generated by their friends. Further, the latter example also shows, in a scale from one to five, how the author of the review has enjoyed a particular location or experience. Conversely, Spotify allows users to make ratings with a dichotomous scale, whereas TripAdvisor does this in a 5-point likert measure. Therefore, at this point, user profiles (i.e. self-presentation) and rating scales (i.e. self-disclosure) start to exert a combined effect on the transmission of information.

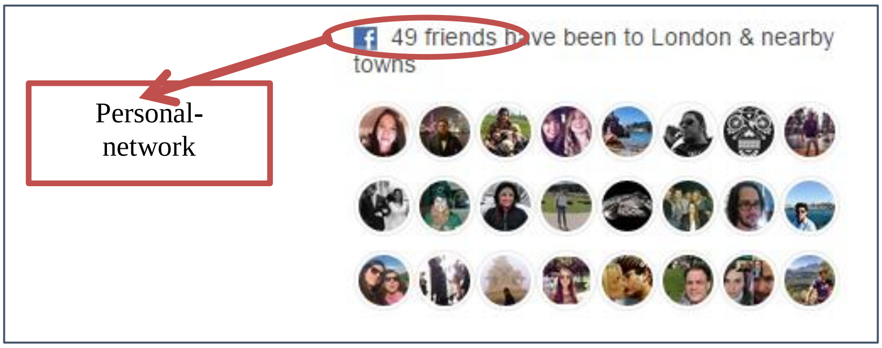


Figure 2.18 – Filtered search by 'friends', restaurants in London (TripAdvisor)

<u>*Conformity to personal networks, differentiated by tie strength (e.g. strong or weak)*</u>

As previously mentioned, there can be a further differentiation among personal networks by introducing Granovetter's (1973) theory of strengths of ties. The two biggest SNSs, Facebook and Google+, allow users to have as many 'smart lists' or 'circles' as they please. However, they do give some suggestions and, in the case of Facebook, it has three predefined classifications of friendship: close friends, friends, and acquaintances (Figure 2.19). Interestingly, Facebook's classification of friendships is similar to how Granovetter (1983) outlined strong, intermediate, and weak ties. Moreover, Facebook also creates (automatic and predefined) lists by individuals' location, place of work, education and family membership.

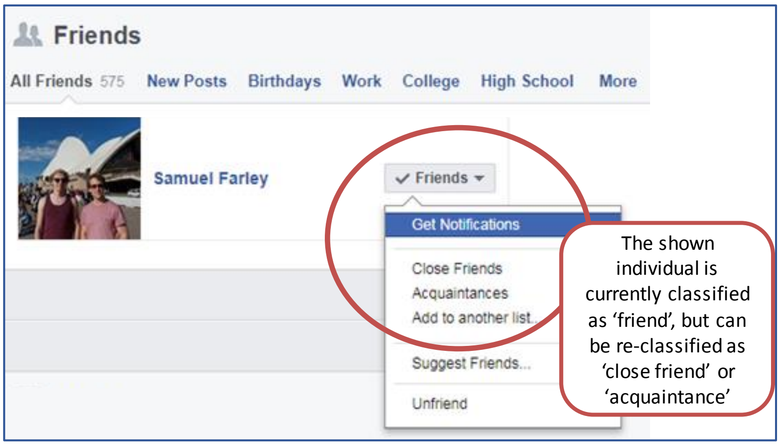


Figure 2.19 - Facebook's pre-defined friendship lists

Regarding the transmission of UGC – which is the primary concern of this thesis –, Facebook's classification of relationships into close friends, friends, and acquaintances means that users can access each of these categories with a click, which makes it fast and easy. Moreover, when an individual creates a post or uploads a picture, (s)he can choose an audience to which this content should be shown. In addition, the broader classification of a person's friends into automatically created lists arguably translates into increased homophily because users would easily access information of those who are similar to them regarding education, job, location, or family. Figure 2.20 shows a list based on the strength of ties (i.e. 'close friends' which is created by default but users need to include people on them) and a 'smart list' (i.e. which is automatically created and updated depending on the personal information shared by users).

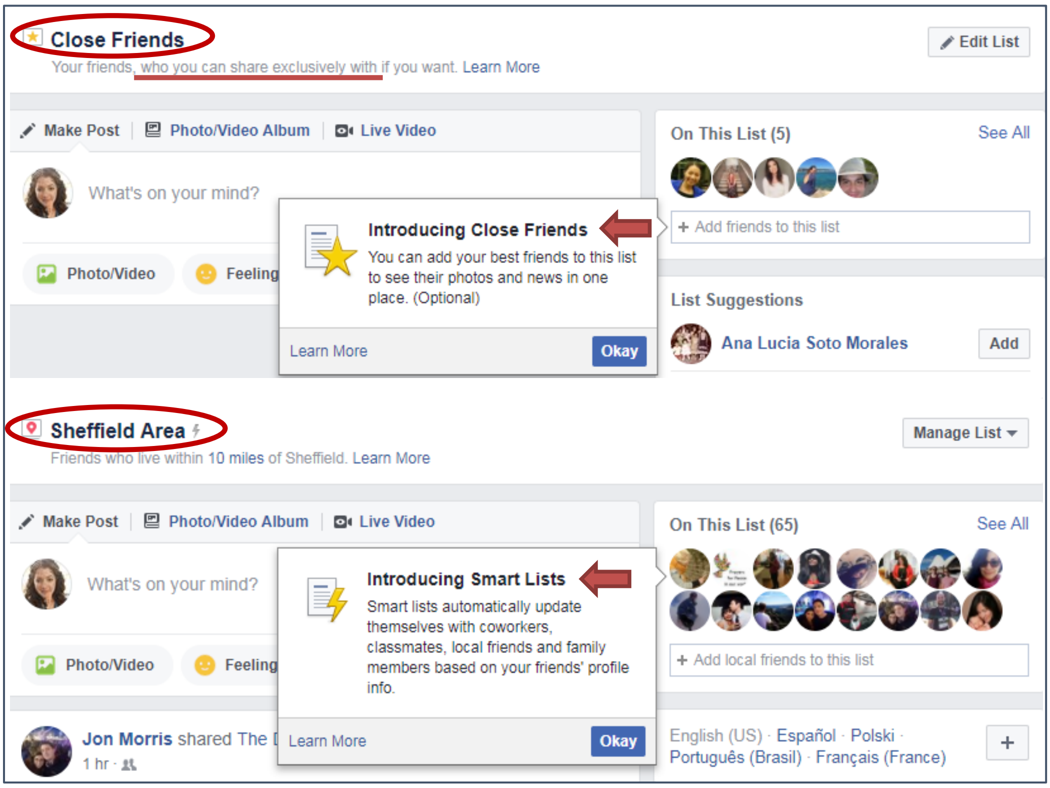


Figure 2.20 - The effect of Facebook's lists on the transmission of UGC

Furthermore, it is not clear to what extent Facebook shares information with other sites (e.g. Wong 2018). However, although it seems that it does not share the detail of its 'smart lists', some sites like TripAdvisor have managed to differentiate strength of ties by highlighting reviews created by Facebook 'friends' (i.e. what could be seen as strong ties) and 'friends of friends' (i.e. which would fall into the definition of weak ties). Figure 2.21 shows how the CC highlights the UGC created by a friend of a determined user, while Figure 2.22 presents how the site makes it easy to access the content of friends of friends. Social network theory shows that if a person A is connected to individuals B and C, the last two are likely to know one another (Granovetter 1973; Jackson 2008). Thus, by having access to the users' real-life ties – which have been declared on SNSs – CCs can exploit the personal networks of users to direct them to their content, arguably, with the effect of biasing their choices. That is, in evolutionary terms, users might be more prone to selecting beliefs from those within their ingroups, and especially from those with whom they have strong ties if CCs exploit the information declared by users in SNSs.

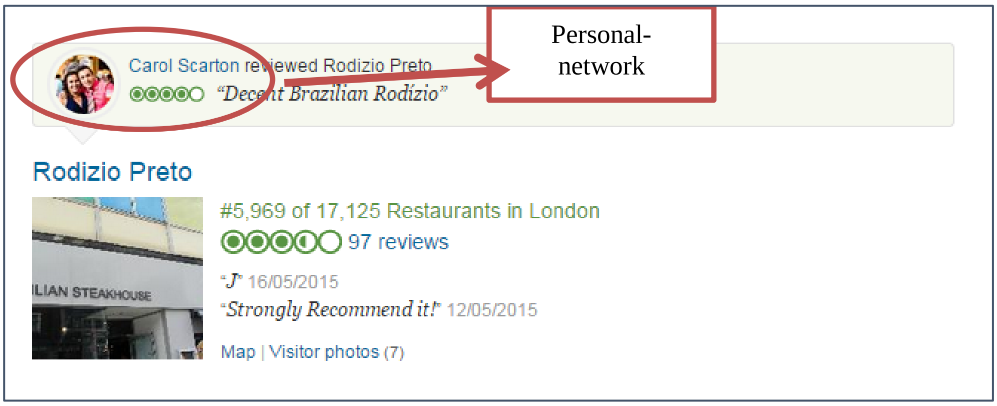


Figure 2.21 – Restaurant review from a (Facebook) 'friend' in TripAdvisor

Figure 2.22 – Restaurant review from a (Facebook) 'friend-of-friend' TripAdvisor

Conclusion

Studies have demonstrated that conformity biases are very powerful in transmitting knowledge (Henrich & Boyd 1998; Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011; Perreault et al. 2012). Moreover, if people tend to imitate individuals who are similar to them and tend to 'follow' and 'friend' like-minded people online (Wu et al. 2011), and add to this that most SNSs only have one-point rating buttons, it becomes quite evident why some of these sites are so attractive for politicians and marketers (Fraser & Dutta 2008). As a consequence, it becomes important to analyse the different ways in which social media sites frame UGC through the designs of their websites, given that these would determine the effect that identity, groups, and relationships may have on the transmission of information and the choices of users.

Lastly, the researcher would like to emphasise that "the adoption of a decision frame is an ethically significant act" (Tversky & Kahneman 1981, p.458). As previously mentioned, individuals can acquire knowledge without being fully conscious of it (Polanyi 1967). Therefore, if social media platforms adopt a frame (i.e. design) that increase the occurrence of biases (i.e. systematic errors) in the process of acquiring knowledge, it is relevant for the academic community and the general public to understand how this might affect the choices of users. As a response to this great need for understanding how the different designs social media sites might enable, restrain, or affect the transmission of information, the following section outlines the theoretical framework used to investigate this process.

2.5 THEORETICAL FRAMEWORK

The fifth and last section of this literature review begins by presenting the thesis' theoretical framework, built from the central concepts and theories discussed so far. Moreover, this section also presents a summary of the outlined literature gaps, together with the aim, research question and objectives of the study.

2.5.1 Theoretical framework: The effect of website designs on the choices of users

The proposed framework builds on the argument that the transmission of UGC in social media occurs through the mechanisms of variation, selection and retention outlined by Campbell (1960), with roots in Darwin's (1859) theory of evolution. Further, this research adopted a Generalised Darwinist position, which allows for the VSR mechanisms to be applied to domains outside biology as long as there is a pool of replicating entities which compete to be selected (Hodgson 2005; Breslin 2010). Thus, for this study, the replicators are people's beliefs, which are then observed through the interactors (i.e. UGC) that users post online. However, through certain elements of their designs, social media sites control to a certain extent the way in which users can present themselves, how much they can disclose, and what type(s) of content they can share (Kaplan & Haenlein 2010). The present thesis focuses on two specific aspects of a website's design: self-presentation, which is seen through the profiles of users; and self-disclosure, which is studied through different rating scales. Both self-presentation and self-disclosure are part of users' identities, which do not always remain constant but instead adapt depending on the presence of groups and relationships (Goffman 1959; Goffman 1963; Tajfel & Turner 1979; Turner 1984; Kietzmann et al. 2012).

From an evolutionary viewpoint, the focal point of this research is the selection mechanism, which can be inferred from the observed ratings people give to the UGC posted online (see Figure 2.5). An argument put forward is that the selection mechanism can be studied through the choice-making process. In this regard, the adopted scheme of prospect theory is based on the assumption that the rationality of individuals is bounded by external and internal constraints, such as the manner in which problems are formulated and people's cognitive limitations (Simon 1955; Simon 1979). Concerning the formulation of a decision problem, prospect theory established that the use of a particular frame has an impact on people's judgement and therefore on their choices (Kahneman & Tversky 1979; Tversky & Kahneman 1981; Kahneman 2003). For the present study, the frame is conceived as the different website designs that might enable, restrain or affect the transmission of UGC.

Moreover, regarding peoples' cognitive limitations, the cost of collecting and examining all information while making decisions necessitates relying on heuristic principles that reduce the difficulty of the task but may lead to systematic errors or biases (Kahneman et al. 1982; Kahneman 2003). Hence, for this thesis, three group-biases that were formulated and have been adopted by a handful of evolutionary scholars were selected: content (i.e. specific qualities of a belief), prestige (i.e. successful individuals), and conformity (i.e. imitating the majority of the group) (Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011). Finally, it is proposed that online environments should differentiate between conforming to the whole network and conformity with users' personal networks, given the outlined effects of groups (Tajfel & Turner 1979; Turner & Reynolds 2012) and the strength of ties (Granovetter 1973; 1983; Cross, Parker, et al. 2001) in the transmission of information.

In order to summarise the theoretical framework of this research show how the different theories interconnect, Figure 2.23 presents the conceptual model of the thesis, highlighting the key theoretical 'building blocks'. As can be seen, the research explores whether certain elements of website designs act as frames and therefore have an impact on the choices of users; that is, on their selection of beliefs, which is inferred through their ratings. In particular, two aspects of the website's designs were investigated: 1) the effect of users profiles (i.e. anonymous or identifiable) on self-presentation; and 2) the impact of rating scales (i.e. likert or dichotomous) on self-disclosure. In addition, because self-presentation and self-disclosure are part of an individual's identity, these will change depending on the presence of groups and relationships, which may be reflected in the choices of individuals. Finally, another aspect that was of interest was to test which of the three group-biases had a stronger effect on choices, and particularly whether people would conform differently to their personal networks when selecting UGC in an online environment.

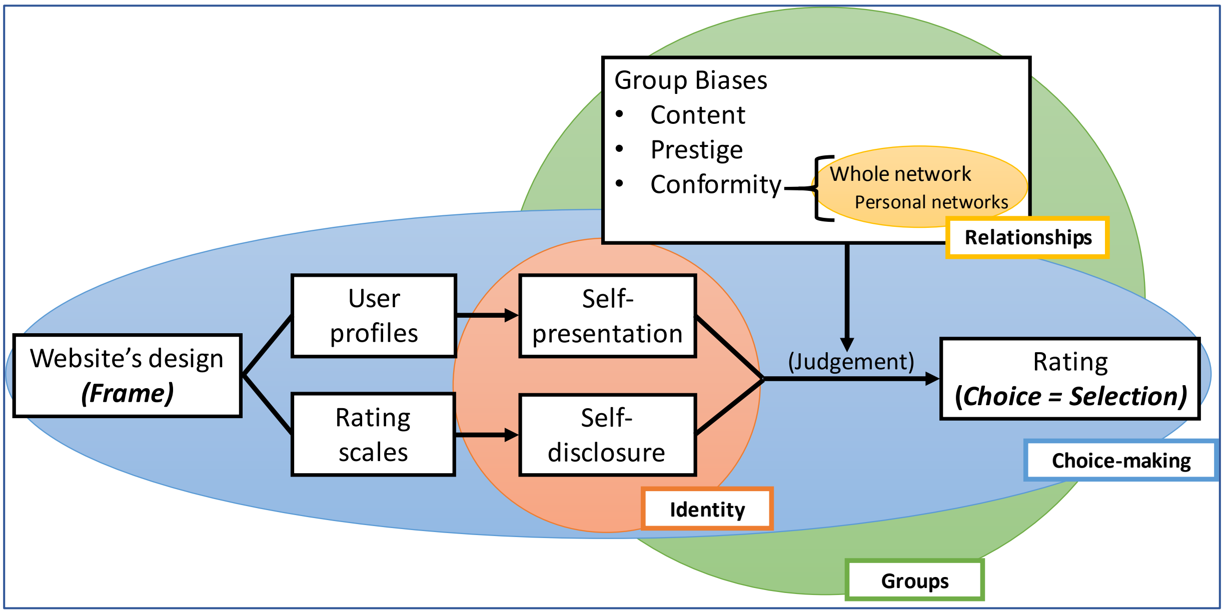


Figure 2.23 - Thesis' conceptual model with the building blocks of social media

It may be argued that the proposed framework presents a certain level of complexity, as there are many concepts to consider. This may be true, but it can also be claimed that Figure 2.23 presents a framework that is more attached to reality. Quoting Simon (1979, p.496):

> "The classical theory of omniscient rationality is strikingly simple and beautiful. Moreover, it allows us to predict (correctly or not) human behavior without stirring out of our armchairs to observe what such behavior is like. All the predictive power comes from characterizing the shape of the environment in which the behavior takes place. The environment, combined with the assumptions of perfect rationality, fully determines the behavior. Behavioral theories of rational choice – theories of bounded rationality– do not have this kind of simplicity. But, by way of compensation, their assumptions about human capabilities are far weaker than those of the classical theory. Thus, they make modest and realistic demands on the knowledge and computational abilities of the human agents, but they also fail to predict that those agents will equate costs and returns at the margin".

2.5.2 Aim & research questions

This thesis aims to determine whether – and how – the different designs of social media websites might enable, restrain or otherwise affect the transmission of UGC, through the impact that these designs have on identity, groups and relationships; and consequently on the choices of users (see figure above). Hence, the overall research question for this project is:

RQ-Overall: How is the transmission of UGC affected by the different designs (i.e. frames) adopted by social media websites?

Empirical evidence is used to address this question throughout the results chapters, but it is specifically tackled in Chapter 7. The rationale behind this query comes, on the one hand, from Tversky & Kahneman's (1981) propositions of how the adoption of frames in a decision problem might lead to people's choices being inconsistent. On the other hand, it comes from the suggestion that online users are prone to searching heuristic content to simplify their choices (Park & Nicolau 2015). However, as social media sites are designed differently, allowing dissimilar levels of self-presentation, self-disclosure, social presence, and media richness (Kaplan & Haenlein 2010); it is likely that these different set-ups might frame information in diverse ways, thus making users more prone to different heuristics and biases that would result in different choices. As previously outlined, this study focuses on two elements of a website's design: user profiles and rating scales; hence, it is centred on social media sites with the same degrees of social presence and media richness, but different levels of self-presentation and self-disclosure. That is, this research imitates and compares the transmission of UGC in a CC and a SNS. It should be highlighted that this study is a pioneer in testing if the different designs of social media frame the content shared online.

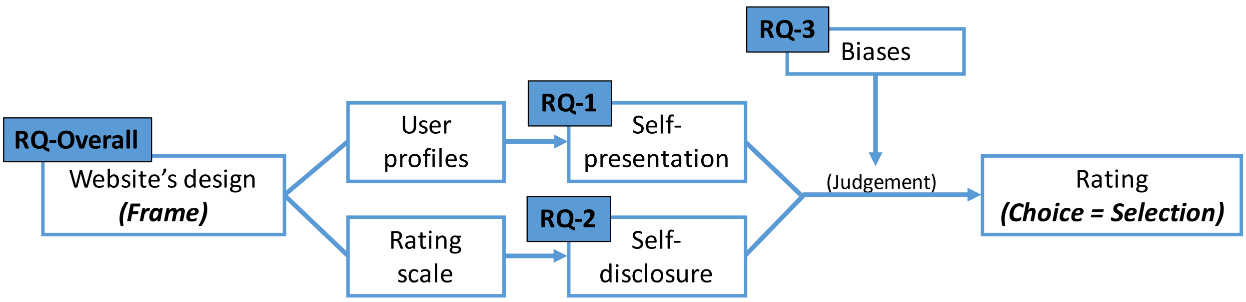


Figure 2.24 - Conceptual model with research questions

Further, the overall question of the study comprises three major research questions, shown in Figure 2.24, and outlined below:

RQ-1: How are the choices of users affected by different levels of self-presentation occurring from diverse user profiles?

The first research question is tackled in Chapter 4. This query is intended to assess only one aspect that differentiates CCs and SNSs: the level of self-presentation (Kaplan & Haenlein 2010). Self-presentation was defined as a component of identity by which an individual tries to control the impression she has on others (Goffman 1959). However, online environments allow users to interact with others without their physical presence, thus allowing disembodiment and anonymity, which allow for distinct ways of identity construction (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008). This research focuses on comparing an anonymous with a non-anonymous environment; the first one being characteristic of CCs while the later one of SNSs (e.g. Figure 2.1 and Figure 2.2). Anonymity has been widely studied, but non-anonymous (i.e.

identifiable) research is scarce (Zhao et al. 2008). Therefore, this research expands the understanding of non-anonymous conditions.

RQ-2: How are the choices of users affected by different levels of self-disclosure happening due to the use of distinct rating scales?

The second research question is addressed in Chapter 5 by the evaluation of another aspect that differentiates CCs and SNSs: the degree of self-disclosure (Kaplan & Haenlein 2010). Self-disclosure was defined as the revelation of personal information that is aligned with the image online users want to give about themselves, such as thoughts, feelings, likes and dislikes (Kaplan & Haenlein 2010). Thus, it was decided to analyse self-disclosure through the comparison of two of the most used rating scales in CCs and SNSs: likert and dichotomous (Table 2.2 and Table 2.3). Regarding theory, it has been argued that rating scales are under-studied and there are no clear guidelines regarding their design (Riedl et al. 2013). Moreover, there is no empirical consensus, as scholars suggest that broader scales are needed (Riedl et al. 2010; 2013) while practitioners favour the use of shorter rating systems (YouTube 2009; Ciancutti 2011). Hence, this study advances current theoretical and empirical understanding of rating scales and their impacts.

RQ-3: How are the choices of users affected by the online presence of their personal networks?

The third research question is tackled in Chapter 6. This question seeks to assess whether online users are more likely to acquire information from those they know offline. There are three reasons that indicate this may be true. First, some evolutionary scholars have found evidence, through models, lab experiments and real occurrences, that individuals are prone to imitating members of their community (Boyd & Richerson 1985; Henrich & Boyd 1998; Henrich 2001; Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011; Perreault et al. 2012). Second, social psychologists argue that individuals' self-esteem is linked to their ingroups (Tajfel & Turner 1979; Turner 1984), which suggests that people might prefer to acquire beliefs from their real-life ingroup if they can differentiate them online. Third, it has been found that the transmission of complex knowledge within networks is favoured by strong ties (Cross, Parker, et al. 2001; Dobson et al. 2013), whereas weak ties are essential for sharing information across the network (Granovetter 1973; 1983). However, despite this evidence, research conducted in online environments has failed to include the effects of conformity, probably due to the complexity that it presents (e.g. Wu et al. 2011; Wilkinson & Thelwall 2012; Liu & Park 2015; Park & Nicolau 2015). Likewise, lab experiments have also avoided measuring the effects of the choices of others (e.g. Riedl et al. 2010; 2013), and the only study that includes social approval cues does not differentiate between the whole and personal network (e.g. Mueller et al. 2018). Therefore, this study is the first attempt to make this distinction in the study of online environments.

2.5.3 Summary of literature gaps and research objectives

Finally, in order to achieve the aim of the study, answer the research questions, and tackle the literature gaps, the objectives of this thesis are outlined in Table 2.4. It should be noted that the objectives are not covered all at once. Rather, they are distributed among the research questions, which correspond to different chapters of results, as outlined above. Chapter 3 includes a visualisation that displays the objectives targeted per chapter of results (see Figure 3.5).

Table 2.4 - Research objectives

Objectives	Literature Gaps
OB-1: To set up an experiment that illustrates how two different types of social media sites (i.e. CCs and SNSs) might enable, restrain, or affect the transmission of UGC.	***LG-1 (p.12):*** There is a need to investigate how the designs of the different classes of social media sites might enable, affect, or restrain the transmission of information.
OB-2: To understand knowledge transmission in social media through the VSR mechanisms; emphasising selection, which is studied through choice-making.	***LG-5 (p.30):*** There is a lack of studies focusing on knowledge transmission with an evolutionary viewpoint. Moreover, this lack of research is more evident in the field of online environments.
OB-3: To investigate if – and how – being identifiable might affect the choices of users, in comparison to being anonymous.	***LG-2 (p.16):*** There is a lack of research exploring the way in which individuals manage their identities when they are known by other members of their network (i.e. non-anonymous).

OB-4: To determine the effects of being identifiable: 1) on real-life interactions, and 2) on the overall attitude towards the website's design.	***LG-3 (p.17):*** There is a need for a greater understanding of how embarrassments take place; specifically, how they affect real-life interactions, and the impact they can have in the design of social media platforms.
OB-5: To investigate if – and how – different rating scales affect the choices of users, by comparing two types: likert and dichotomous. **OB-6:** To determine if the use of different scales has: 1) any social implications when users are identifiable, and 2) any effect on the attitude towards the website's design.	***LG-4 (p.21):*** There is a theoretical and practical need to deepen the understanding of rating scales. And, arguably, the study of ratings should be done within the context of users' identities, groups and relationships.
OB-7: To detect how the three group-biases take place in online environments, focusing on the study of conformity, which has received limited attention both offline and online.	***LG-6 (p.43):*** There is a gap in studying imitation heuristics and biases with a Darwinian perspective. What is more, conformist transmission remains poorly studied. ***LG-7 (p.44):*** Studies perform in CCs and SNSs have studied what could be catalogued as content and prestige-based biases, but only one has been found to study conformity.
OB-8: To determine if – and how – users' real-life ingroups and outgroups affect their choices, and thus the transmission of UGC.	***LG-8 (p.47):*** There are theoretical and empirical gaps that need to be addressed about the presence of groups in online environments. Moreover, regarding research, the only experiment that studies conformity does not differentiate between that to the ingroup and outgroup.
OB-9: To make a further differentiation of ingroups using the strength of ties, and evaluate if these have any effect on choices. Further, to assess if online interactions can predict real-life relationships. **OB-10:** To understand: 1) the effect that online interactions might have on real-life relationships and vice-versa, 2) the extent to which online users are aware of the impact that others have on their choices.	***LG-9 (p.50):*** Regarding the strength of ties, there is a need to join social media predictions with real-life data. ***LG-10 (p.51):*** There is a necessity to describe how can online interactions affect offline relationships and vice-versa.

CHAPTER 3: METHODOLOGY

The aim of this chapter is to elaborate on how the thesis' conceptual model is to be studied. The chapter is structured as follows: first, the philosophical underpinnings of this research are outlined. Second, the current methodologies used in social media research are discussed, highlighting their advantages and shortcomings. Third, the experimental set-up is described, together with the chosen methods for data collection, justifying these decisions with the philosophical position and the discussion of methodologies used in the field of social media. Fourth, the way in which the data were collected is explained, emphasising the sources of data, their limitations and ethical considerations. Finally, the process of data analysis is presented, linking it with the conceptual model and the objectives of this thesis.

3.1 PHILOSOPHICAL VIEWPOINT AND ASSUMPTIONS

It is of relevance to outline the philosophical paradigms embraced in this study, as these have implications regarding the nature of the social world (i.e. ontology) and the type of knowledge that can be acquired from it (i.e. epistemology) (Johnson & Duberley 2000). What is more, adopting a specific paradigm depends on the conceptions of the problem that is to be investigated (Gill & Johnson 2010; Bryman & Bell 2011). In the case of this thesis, its main aim is to determine whether – and how – the designs of social media sites might enable, restrain or otherwise affect the transmission of UGC, through the impact that different set-ups have on identity, groups and relationships; and consequently, on the choices of users. Thus, the examination of this issue requires investigating the causality of different elements on the choices of users. In addition, the response to the posed aim is expected to be generalisable to a certain extent. As a consequence, both causality and generalisation assumptions locate this research within positivist approaches (Eisenhardt 1989).

The current research adopts a post-positivist standpoint. The implications of adopting this position are that, from an ontological perspective, it is believed that there is an independent reality, very much like positivism. However, despite the shared assumption of ontological realism, the posed research questions and objectives are meant to be addressed in line with Popper's Falsification principle (1959), rather than following the logical positivist idea of verifiability. Further, post-positivists also acknowledge that reality can only be known imperfectly and probabilistically (Robson 2011).

Moreover, from an epistemological viewpoint, the search for knowledge is centred on causal explanations from observed patterns within the social world. Therefore, knowledge is thought of as being objectively obtained (Miller 2005). However, unlike classical and logical positivists, post-positivists acknowledge that everyone is culturally biased and hence individuals cannot achieve objectivity in a perfect manner, but merely approach it. For this reason, researchers are encouraged to obtain multiple observations from different sources, so as to triangulate across multiple imperfect perspectives (Trochim et al. 2016).

To sum up, in line with post-positivism, the present thesis assumes a modified objectivist epistemology and accepts that reality exists but is only 'probabilistically apprehendable' (Guba & Lincoln 1994). These assumptions have a number of implications for this study. Firstly, regarding the conceptual model of the research (see Figure 2.24), it is believed that the effect of website designs on users' choices can be measured. Therefore, patterns concerning the effects that sites have on the transmission of UGC can be discovered and generalised to a certain extent. Secondly, to objectively understand this issue, the researcher must be able to triangulate different perspectives. Hence, more than one source of data was utilised. Thirdly, it is acknowledged that regardless of the chosen methodologies to collect and analyse data, these still have limitations in capturing an objective reality. For this reason, Section 3.3.4 outlines some of the limitations concerning the chosen methods of data collection.

3.2 METHODS

Research methods denote the techniques and procedures that researchers use to obtain information to address the research questions; they comprise the study design, data collection, and data analysis (Bryman 2012). Methods are usually catalogued as *quantitative* when the quantification in the collection and analysis of data is emphasised, or *qualitative* when greater importance is given to narratives (Bryman & Bell 2011).

Further, it has been argued that the choice of research method should be aligned with the adopted philosophical stance (Burrell & Morgan 1979). Thus, to select the method – or combination of methods – that would better support this research, the current approaches used in social media studies were assessed regarding their strengths and limitations. Afterwards, also taking into account the adopted philosophical viewpoint, the methods for this study were chosen. Hence, the following sub-section outlines an assessment of the methods utilised in social media research, succeeded by the ones adopted in this thesis, and their rationale.

3.2.1 Methods used in previous social media research

1. ***Extracting available data from social media:*** A large number of researchers extract UGC from what could be called its 'natural environment' (i.e. online). Subsequently, most of them apply *social media analytics* to perform the analyses. That is, they make use of quantitative methods such as sentiment, content, and word-frequency analyses (e.g. Wu et al. 2011; Wilkinson & Thelwall 2012; Procter et al. 2013; Thelwall et al. 2012; Cheng & Ho 2015; Park & Nicolau 2015; Liu & Park 2015; Goel & Goldstein 2014; Godes & Mayzlin 2004; Jacobsen 2015). It should be noted that a small number of academics instead make use of qualitative methodologies, such as online ethnographies or '*netnographies*' (e.g. Lantz-Andersson et al. 2013; Kozinets 2002). However, qualitative methods seem to be less common because of the nature of collected data, which tend to be extracted in huge volumes. Some of the advantages of using the internet for data collection are that it is time and money efficient, geographic location is not an issue, and large volumes of data can be easily obtained and analysed (Bryman & Bell 2011). On the other hand, as the internet is not yet accessed by everyone (e.g. there are geographical, political, and age limitations), some sectors of the population are not represented in the research. Therefore, generalisations can be misleading. Moreover, there is arguably a loss of 'personal touch' given that in most cases this data collection method does not allow researchers to know the identities of those whom they analyse. And, most importantly, there are ongoing ethical debates concerning the lack of consent from the people whose data are extracted and later used for research purposes (Bryman & Bell 2011). A final

limitation of this method of data collection is that most researchers use the social media platforms with fewer restrictions on the extraction of data (i.e. Twitter), which reduces the options for comparison between platforms.

2. ***Setting up an experiment in an existing social networking site or content community:*** Some researchers have set up controlled experiments using existing SNSs or CCs (e.g. Kump et al. 2013; Schweisberger et al. 2014; Tsao et al. 2015). These kinds of experiments possess the same advantages of those where data is extracted from social media. Also, as individuals willingly take part in the study, the researcher is privy to the identity of its audience and the connections among them, so ends up with a “captive population who are already communicating with each other” (Bryman & Bell 2011, p.656). However, this method has the limitation that there could be less spontaneity of response as individuals are more likely to review their words before they post them online (Bryman & Bell 2011).

3. ***Creating a web-based experiment:*** An increasing number of researchers are making use of web-based experiments, by simulating an online community (e.g. Eryarsoy & Piramuthu 2014; Lee et al. 2012; Riedl et al. 2013; Riedl et al. 2010; Mueller et al. 2018). As discussed above, in these situations the researchers have the advantage of knowing their studied population. Moreover, a further advantage of this method is that users are less exposed to external factors that can affect their choices, such as advertising or suggestions from the search engines. However, with these experiments, some of the ‘online reality’ is lost. For instance, some researchers set up the experiments in such a way that they prevent users from being biased by the comments of other participants (e.g. Riedl et al. 2010; Riedl et al. 2013), arguably losing essential features of social media communication.

4. ***Modelling online social interactions:*** Some researchers have modelled different aspects of social media (e.g. Siegel 2013; Jiang et al. 2014; Gopinath et al. 2014; Sun et al. 2015; Jiang et al. 2015; Riedl et al. 2013). The benefit of using this method is that a model enables reality to be simplified, therefore facilitating decision-making, control of variables, and predictions. However, a shortcoming is that, because of its simple nature, it does not comprise all aspects that affect what is being modelled (Thiétart 2001).

5. ***Making use of surveys and interviews:*** Methods such as online surveys and interviews are widely used to obtain a higher level of detail regarding users' online preferences. Researchers who use online surveys (e.g. Chen & Liang 2011; Correa et al. 2010; Hughes et al. 2012; Lee et al. 2012; Lee 2014) have the advantage of being able to reach many potential interviewees at low cost and effort. However, they also tend to have modest response rates, and samples can be considered biased due to the nature of internet-user populations (Bryman & Bell 2011). Conversely, researchers who make use of interviews (e.g. Quan-Haase & Young 2010) have the benefit of greater flexibility and, as they get to see the interviewees, gain more insights from their gestures and expressions. Nevertheless, data gathered from interviews are more time and effort-consuming, and it is harder to generalise insights (Bryman & Bell 2011). Also, it could be argued that only using surveys or interviews without accompanying them with any of the above-listed methods makes little sense, as the nature of social media is precisely the online interaction itself.

3.2.2 Proposed methods for data collection

To address the research questions of this thesis, and in an attempt to overcome some of the limitations of previous studies, the present thesis opts for a mixed and multiple methods approach. The difference between a multi- and a mixed-methods approach is that the former concerns the use of multiple data sources within one methodological paradigm. Conversely, in the latter, a different type of information – quantitative and qualitative – is to be obtained (Denzin & Lincoln 2018).

It has been argued that the use of multiple methods helps attain more profound insights and more reliable results (Boudreau et al. 2001; Riedl et al. 2010). Specifically, research adopting a multi-method approach has provided some of the most reliable results on the topic of collective decision-making within online environments (Riedl et al. 2010; 2013). Similarly, it has been claimed that the use of mixed-methods brings a significant advantage as it enables the clarification and more in-depth knowledge of particular aspects that are detected as being important (Bryman & Bell 2011). Notably, some researchers have stressed the need for mixed-methodologies to gain a better understanding of identity construction in online environments, arguing that "the next logical step in advancing this line of research is to combine investigators' objective coding of the profiles with users' subjective interpretations of their own activities" (Zhao et al. 2008, p.1832). Further, a number of scholars studying the designs of websites advocate for the use of mixed-methods as they prove to be "effective to attain a deeper and more comprehensive understanding" (Cyr et al. 2009, p.558).

However, from a philosophical viewpoint, mixing qualitative and quantitative data can give rise to specific issues. On the one hand, some researchers claim that combining these methods involves different epistemological perspectives, and it is therefore wrong to use them together (e.g. Guba & Lincoln 1994). On the other hand, scholars in favour of mixed-methods argue that philosophical positions should not dictate or interfere regarding which methods are to be used to gain knowledge (e.g. Johnson & Onwuegbuzie 2004). Instead, such academics recommend that the research questions, and not the philosophical positions, should define the methods to be employed.

In the case of this study, mixed-methods have been chosen because the adopted philosophical stance – post-positivism – encourages researchers to obtain multiple observations from a range of sources in order to triangulate across multiple imperfect perspectives (Trochim et al. 2016). Nonetheless, the research questions and objectives have also influenced this decision. For instance, the first objective of the study was set to determine how the choices of users are affected by different user profiles, alternative rating scales, and the presence of their personal networks. Hence, to make this determination, the data should be as objective as possible, and should therefore come from a quantitative source. Conversely, to deepen the understanding of the effect that identity, groups and relationships have on choices, it is better for the data to be of qualitative provenance. Therefore, in line with the adopted philosophical position and methodology, the present study makes use of three primary sources of data: online interactions arising from a quasi-experiment, questionnaires, and focus groups. Section 3.3 details how the experiment was designed and the manner in which the data were collected, and Section 3.4 describes the methods and procedures for the analysis of data.

3.3 QUASI-EXPERIMENT SET-UP

This research had a number of particulars that required it to be designed as a quasi-experiment. To start with, the study needed to allow the comparison of key characteristics of SNSs and CCs. However, this could not have been possible if data were merely extracted from these sites because the information would not be comparable. For instance, if the researcher had 'crawled' data from Facebook and TripAdvisor it would have been impossible to determine the effect of the website designs on ratings, given that the content shared on these sites differs significantly. As a consequence, it was necessary to set-up an experiment. However, this experiment needed to be as 'real' as possible, and thus it required an existing SNS or CC. Also, as outlined in Chapter 2, this research proposed to differentiate between the whole and personal networks, and among different strengths of ties. Hence, setting-up a one-day experiment would not have been sufficient as people taking part in the experiment needed to have real-life friendships and interact over a period of time. Ideally, participants had to interact over a more extended period in a platform where some users were unknown to them, whereas others were real-life friends.

To fulfil the conditions that needed to be in place for the objectives to be attained, the study made use of an existing CC with users having some of their real-life friends also using the site. However, the reality gained meant having less control over certain elements. When the researcher has no authority on the assignment of participants to experimental and control situations, these are called 'quasi-experiments' (Campbell & Ross 1968). The current study, therefore, was set-up as a quasi-experiment. These are characterised by having non-randomly selected groups and are usually longitudinal studies involving a pre-test, an experimental treatment and a post-test (Campbell & Ross 1968). However, a limitation is that in the social sciences researchers can rarely eliminate 'disturbing influences' and "must rely on evidence cast up by the 'experiments' that happen to occur" (Friedman 1953, p.150). Nevertheless, this is not necessarily considered to be a shortcoming, as it has been argued that no experiment can be completely controlled (Friedman 1953). Therefore, it is critical for researchers to understand and explain the limitations of the simplified model and the alternative explanations (Campbell & Ross 1968; Donaldson 1997; Wall et al. 1986). To validate the findings and be able to offer alternative explanations, the triangulation technique was used. As outlined, and in accordance with the philosophical position and methodology adopted, triangulation requires the use of more than one method on the study of social phenomena and is widely used with mixed-methods, as these can help both achieving convergent validity and completeness (Yu 2005; Bryman & Bell 2011).

3.3.1 Research context and overview

As mentioned, it was necessary for this study to use the same online platform so that information could be analogous. However, at the same time, it was crucial that such a platform allowed the comparison between a SNS and a CC. Thus, a required feature for the online platform was the possibility of modifying both user profiles (anonymous versus identifiable) and rating scales (likert versus dichotomous). To fulfil these conditions, the researcher chose an existing CC and gradually implemented features of a SNS.

The selected site was *PeerWise (https://peerwise.cs.auckland.ac.nz)*, which is a free educational platform that students use to "create [multiple choice questions] and to explain their understanding of course-related assessment questions, and to answer and discuss questions created by their peers" (PeerWise 2015). The main reason why this CC was chosen was that the creator of PeerWise agreed to make changes on the website in order to suit the needs of this research[14]. Moreover, PeerWise was already being used in a module at the University of Sheffield, which facilitated access and allowed the researcher to analyse data from a previous year. These situations created an ideal environment that made possible a natural, longitudinal, quasi-experiment.

The quasi-experiment consisted of three years of data generated in the online educational environment where the participants were non-randomly allocated final-year undergraduate students that took a designated (core) module/course at the University of Sheffield. The module used PeerWise as an ongoing assessment throughout the semester, where students used it to post, answer, rate, and comment on multiple choice questions related to their module. Hence, the data consists of three years of online interactions, comprised in the three quasi-experimental designs of each cohort. Figure 3.1 presents a visual representation of the quasi-experimental design.

[14] Changes on PeerWise only affected the module in the University of Sheffield that was used for the experiment. All other modules from this and other universities were unaffected.

As can be seen in the figure below, the first stage focused on altering the levels of self-presentation through the comparison of anonymous and identifiable users. The first stage tackled the first research question, which examined how the choices (i.e. ratings) of users were affected by different levels of self-presentation (Chapter 4). The second stage was centred on the modification of the levels of self-presentation, through the comparison of likert and dichotomous rating scales. The second stage tackled the second research question, which investigated how the choices of users were affected by different levels of self-disclosure (Chapter 5). Moreover, the third research question – which focused on the study of personal networks – was also investigated with the data generated in the second stage, as users were all identifiable (Chapter 6). Finally, the overall research question – which was set to understand how the transmission of UGC and the choices of users were affected by different designs of social media sites – made a comparison of the three quasi-experimental conditions (Chapter 7).

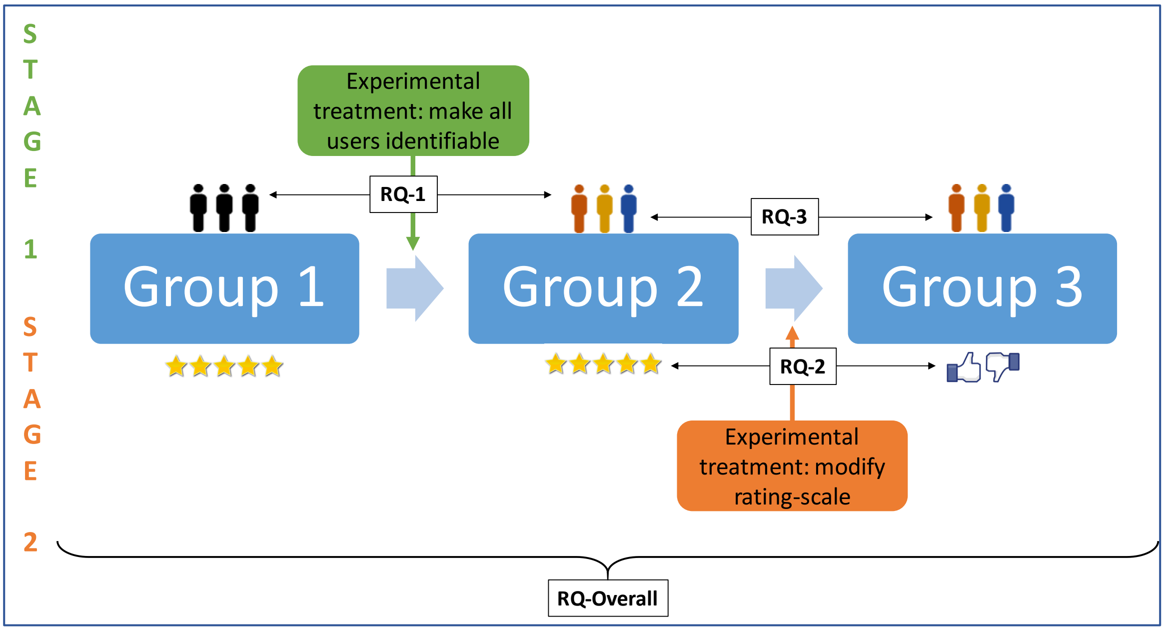


Figure 3.1 - Quasi-experimental set-up.

Regarding each cohort, Group 1 comprised the online interactions (i.e. posted questions, answers, ratings, comments) generated in the Autumn semester[15] 2014-15. This group was presented with the original design of PeerWise, where users interacted

[15] Autumn semesters in the UK usually start in September and finish in February of the following year.

anonymously[16] and could rate the posted questions using a 0-to-5 likert scale. Group 2 comprised the data from the Autumn semester 2015-16. Here, users continued to use the likert scale but were made identifiable, by signing in with their real identities (i.e. *Name_LastName*) which were pre-populated by administrators based on the list of students participating in the module. Finally, Group 3 comprised the interactions from the 2016-17 Autumn semester, where users continued to be identifiable while the rating scale changed from likert to dichotomous (like-dislike). It should be highlighted that, regarding self-presentation and self-disclosure, the way in which PeerWise was originally set-up resembled a CC (Group 1), while by the end of the quasi-experiment the adopted design imitated a SNS (Group 3).

It should further be highlighted that Figure 3.1 presents a very simplified version of the experimental changes implemented. In reality, PeerWise had to undergo a number of adaptations to 1) make users not only identifiable, but also aware of how others were answering, rating, and commenting to others; and 2) modify the rating scale in such a way that allowed comparison between the studies and did not affect other indicators on the system, such as badges and leader-boards. Appendix 3.1 presents the full documentation of changes performed within PeerWise[17]. It is recommended that the reader consults this to get a clearer idea of: how PeerWise is set-up, the quasi-experimental design, and the type of online interactions. Moreover, each chapter of results (Chapters 4-7) contains a brief methods section which includes a picture of the experimental treatments in PeerWise (see Figure 3.1).

3.3.2 Sample and timeline for data collection

The population for the quasi-experiments consisted of 973 third-year undergraduate students that took the same module, in the same University, with the same module leader and teaching staff. It should be noted that, in all three cohorts, about three-quarters of the group was based in the UK, while the other quarter was based in China (but with access to the same module materials including lectures). Moreover, from the University of

[16] Users were able to choose their usernames when they first sign in. However, PeerWise is designed so that all content appears anonymously.
[17] All changes were designed by the researcher (Gabriela Morales), and these were later agreed with and implemented by the creator of PeerWise (Dr. Paul Denny, from the University of Auckland, NZ). All changes were also approved by the *Ethics Committee of the University of Sheffield's Management School* and by the module leader (Dr. Jon Burchell, from the University of Sheffield).

Sheffield's database it was obtained that students were from 56 different nationalities, although the majority were British (48.7%), followed by Chinese (28.1%). Also, 53.1% of the sample were females, while 46.9% were males. Further, from the surveys (see below), the average declared age was 21.4 years of age.

At this juncture it is appropriate to describe briefly the timeline for the data collection, as this had an impact on the way in which students interacted on PeerWise, and possibly on the data collection. As mentioned above, each group of data collection corresponds to an Autumn semester of a module taught at the University of Sheffield. Hence, Figure 3.2 presents how the three sources of data (online interactions, questionnaires, and focus groups) were collected, in relation to the weeks of the semester. Moreover, it highlights when the use of PeerWise was mandatory, as opposed as when it was optional. It should be noted that questionnaires and focus groups were only collected for Groups 2 and 3, for reasons that will be further explained within this chapter.

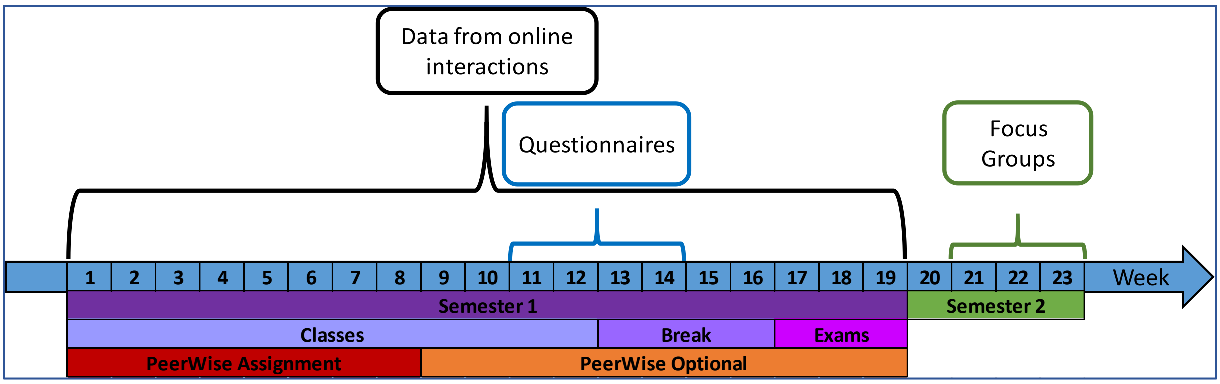


Figure 3.2 - Timeline for data collection

As can be seen in the figure above, data from the online interactions were generated and collected through the 19 weeks that the semester lasts. However, although PeerWise was available throughout the semester, the assignment which mandated its use only lasted eight weeks; after this period, the use of PeerWise became optional. However, it was discovered that several students continued to use the site beyond the required period, using it as a revision tool for their final exam. Further, the questionnaires were conducted after the assignment took place but before the students received the marks, in order to avoid responses being affected by grades. Finally, due to time constraints during the exam period, focus groups had to take place after the semester was over and once students had received their marks for the assignment, but before they received their final grades (incorporating the exam component).

3.3.3 Data collection

As previously outlined, data were collected from three main sources: 1) online interactions that were extracted from each quasi-experimental set-up, 2) questionnaires, and 3) focus groups. Table 3.1 presents a summary of the data collected among the three groups. It should be noted that the first cohort used PeerWise while the quasi-experiment was being conceived. Therefore, this group contains only the retrieved online interactions, as it was not possible to conduct questionnaires or focus groups. Moreover, it should be stressed that the online interactions listed in the table below are only the main ones; additional interactions will be described in the following sub-sections, which explain in detail each of the sources of data.

Table 3.1 - Summary of collected data

	Group 1	Group 2	Group 3	Total
Participants (active on PeerWise)	295	369	309	**973**
Online interactions	74,463	76,685	41,922	**193,070**
- Posted Questions	*2,104*	*2,525*	*2,125*	***6,754***
- Answers	*34,018*	*34,379*	*21,486*	***89,883***
- Ratings	*25,927*	*25,976*	*14,997*	***66,900***
- Comments (total)	*12,414*	*13,805*	*3,314*	***29,533***
Questionnaires	*N/A*	186	206	**392**
Focus groups	*N/A*	x3 (17 students)	x3 (12 students)	**x6 (29 students)**

1. ***Data from online interactions***

The primary source of data collection corresponds to online interactions. Setting up an experiment on a social media site has the advantage of attaining online reality while having a ‘captive population’ where users are already communicating with one another (Bryman & Bell 2011). Further, in contrast with interviewing participants, analysing online interactions helped to uncover patterns of which they might themselves be unaware.

As shown in Figure 3.2, data from PeerWise comprised 19 weeks of interactions for each year of the quasi-experiment. Specifically, for the eight weeks that the assignment took place, students were required to engage with the educational website by: 1) authoring a minimum of five multiple-choice questions, 2) answering at least 20 questions from their peers, and 3) engaging with PeerWise throughout the assignment. Ratings and comments were optional and were not considered for their assignment. These requirements for the assignment remained unchanged[18] for the three years that the study lasted.

It should be noted that the posted questions were non-cumulative. That is, each group saw only its own questions. However, in order to increase reliability and to be able to make an 'equal' comparison between groups, a few questions were identical over the three groups (see Figure 3.3). These questions were posted in PeerWise at the start of each academic year, appearing under the authorship of the module's staff.

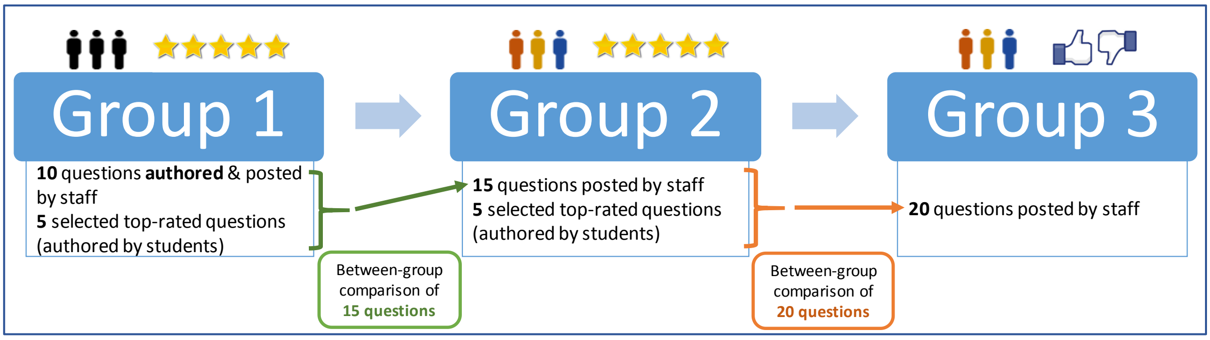


Figure 3.3 - Identical questions over the three cohorts

Furthermore, as can be seen in Appendix 3.1 – where all the screenshots of PeerWise are shown – the website allowed for more interaction than only questions, answers, ratings and comments. For instance, when authoring a question, students could add a tag or explanation. Further, after answering a peer's question, students had the option to 'follow' the author (similar to the 'follow' function on Twitter). Moreover, to incentivise students to engage with the website, PeerWise adopts various gamification principles and gives

[18] The assignment within PeerWise remained unchanged for the three years when data were collected (i.e. authoring five questions, answering 20 questions from their peers, and showing continuous engagement through the duration of the assignment). However, outside PeerWise, Groups 1 and 2 students were assessed by presenting what they considered their three best questions, whereas in Group 3 students needed to write a reflective essay regarding their best question. Given the engagement elements of the assignment remained unchanged, this was not considered to have had an effect on the interactions within the website.

users 'badges' which are then translated into a 'reputation' score. The students with the highest scores appear on leaderboards.

Online interactions were extracted from the website and served to explain, quantitatively, the effect that different variables had on the ratings given by students. Therefore, the most crucial variable was ratings, as these represented the evaluations (i.e. choices) of users. Figure 3.4 shows the way in which the collected data relate to the proposed conceptual model and the research questions. It should be noted that not all variables were tested at once. As Figure 3.1 showed, the first stage of the quasi-experiment focused on comparing two designs of user profiles, while the second stage compared two rating scales. Thus, only one change was made and tested at a time.

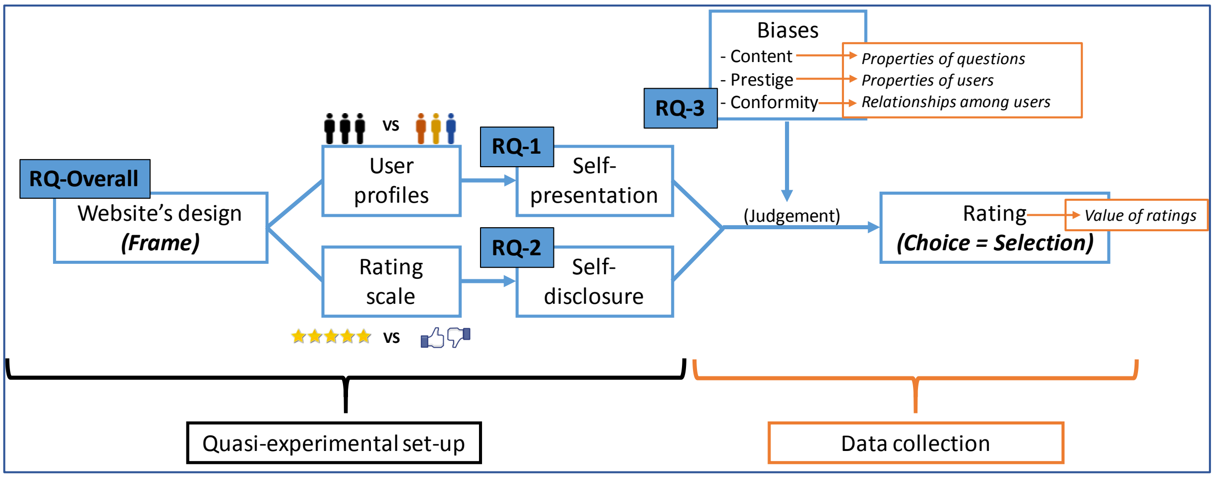


Figure 3.4 - Linking the conceptual model with the data collection

Ratings were the dependent variable. Thus, the *values of ratings* were analysed individually and as averages. Moreover, the independent variables were comprised in three groups (see below), corresponding to the three group-biases described in Chapter 2; content, prestige and conformity. It should be noted that a full list and description of all individual variables are presented at the start of Chapters 4 and 5.

- *Properties of questions*: These online interactions, extracted from PeerWise, were catalogued as content-based biases (i.e. specific qualities of a belief). Examples of these variables are the length of the posted question, its readability index[19], whether it included a reference or a link to a video or article, and so forth.

[19] A test designed to evaluate how readable a text is

- *Properties of users (who authored the question)*: These interactions were considered to reflect prestige-based biases (i.e. successful individuals). Some of these variables were used to measure 'online gained prestige', like number of badges and reputation scores from PeerWise. Moreover, other characteristics helped to reflect 'real-life prestige', such as gender and nationality, were obtained from the University's database.
- *Relationship among users (between author and 'rater[20]')*: These interactions were thought of as reflecting conformity-based biases (i.e. imitating the majority of the group). As proposed in Chapter 2, this thesis differentiates among the whole and personal networks, and then according to the strength of ties (i.e. strong, intermediate, and weak). This information was collected with the use of the surveys and will be explained in the following section.

It should be highlighted that students were not directly assessed on PeerWise, but were rather asked to present what they considered to be their three best-authored questions, or to reflect on their best question in the final cohort. Later, these questions were marked by the module staff. It was thus clear from the start that any score they would receive within the educational site (e.g. ratings from their peers, correct or incorrect answers, badges, or their appearance on the leader-boards) would *not* affect their grades in any way. Although nothing from the experiment would affect the grades of students, all were asked to consent to take part in the experiment the first time they logged into PeerWise. If a student did not wish to take part, their authored questions were excluded from the analysis and they were not emailed with an invitation to answer the survey. As for the online interactions of Group 1, only aggregated, anonymised data were retrieved. Moreover, these interactions were later matched with the students' demographics by a departmental administrator, so the researcher never had access to the names or emails of this cohort. Finally, the *Ethics Committee of the University of Sheffield's Management School* approved both the quasi-experiment design and the extraction of data.

2. *Online self-completed questionnaires*

The purpose of questioning participants is to understand their "own behaviour or that of others, attitudes, norms, beliefs and values" (Bryman & Bell 2011, p.201). Specifically,

[20] The word 'rater' was used to denote the user who rated a question posted by the 'author'

self-completion questionnaires fulfil this function but are different from interviews in the way that they have fewer open-ended questions, are shorter, and most importantly, participants go through the questions themselves. Some advantages of self-completed questionnaires – and why they were considered appropriate for this research – are that they are quicker to administer, diminish interviewer effects and variability, and are more convenient for respondents who can complete them at their own pace (Bryman & Bell 2011).

However, there are a few disadvantages of using questionnaires. First, respondents cannot ask questions if they need clarification (which may be crucial for those who are not fluent in English). Second, there is no opportunity to question participants further/probe responses. Third, respondents are more prone to fatigue, which may lead them to abandon the questionnaire. Fourth, in some cases, participants can read the whole questionnaire before answering the first question. Fifth, it is argued that the researcher can never be entirely sure of who has responded to the questionnaire; and sixth and finally, these are usually characterised by more missing data and lower response rates (Bryman & Bell 2011).

Consequently, to mitigate some of these shortcomings, the questionnaires were developed as follows. First, the questionnaire was designed in a way that took between 7 and 15 minutes to respond, to avoid response fatigue. Second, all questions were written using simple English words, considering that a high percentage of students were not English native speakers. In addition, the questionnaire was checked for clarity of language and errors by the researcher's supervisory team; and for usability on different mobile devices by one of the supervisors and also other PhD students. Third, questions that asked for personal information displayed an explanation of why this data was needed and how it would be used, to prevent students from abandoning the questionnaire. Fourth, the software used to conduct the survey (*Qualtrics*), had a number of built-in options to minimise the deficiencies of online self-completed questionnaires. For instance, participants received an invitation to answer the survey containing a personalised link, which allowed them to save their progress, so they did not need to start from the beginning if they closed the query. Further, having individual links ensured that there was only one response per person. Fifth, respondents could not go back to preceding sections or move forward without completing all mandatory questions of a section. This was implemented to avoid users from reading the entire questionnaire before answering it or going back to change their response when realising there was a follow-up question. Sixth, Qualtrics allowed questions to be mandatory or optional. The questionnaire was set up so all close-ended questions were compulsory while all open-ended ones were elective. This made the questionnaire quicker while at the same time allowing participants to elaborate on the answers they considered necessary. Finally, the option ‘randomisation of questions’ was used to avoid response order effect, while ‘automate the sequence of questions’ was used so participants were only presented with those queries they were required to answer (the latter option is mostly used for follow-ups).

As explained, questionnaires were distributed and completed while students were still using PeerWise, but before they received the grades for their assignments, to avoid biases regarding their marks. Moreover, to minimise missing data and prevent low response rates, a raffle of Amazon vouchers was conducted. The raffle was approved by the Ethics Committee of the University of Sheffield’s Management School and involved one £50 voucher and five £20 vouchers (£150 in total, per cohort). In addition, it should be mentioned that questionnaires were optional and were only sent to those students who agreed to their online interactions being used in the study. Appendix 3.2 shows all questions included in the questionnaire.

The questionnaire had two purposes. Firstly, it was aimed at understanding how students felt about being and having their peers identifiable, and their perceptions on the use of the two tested rating scales. This aim was achieved by collecting quantitative data through close-ended questions, and qualitative insights through open-ended questions. Secondly, the questionnaires helped to unravel the relationships among users. Making use of the theory of strength of ties (Granovetter 1973; 1983), and based on how this concept has been applied in the study of social networks (e.g. Jackson 2008), the questionnaire asked students to first name at least three people they knew within the module. Then, in a subsequent question, they were asked how often they saw their peers outside the class, and this information was used to determine the strength of the ties. A more detailed information on this will be presented on Chapter 6, which is dedicated to the study of personal networks. Once the strength of ties was known, these were used to determine if friendships had any effect on ratings.

Finally, a summary of the number of responded questionnaires is displayed in Table 3.2. Although there is a debate on the ideal response rate, for this research 50% was considered adequate, 60% good, and more than 70% very good (Rubin 2010). As can be seen for groups 2 and 3 the response rates were of 52.0% and 70.1%, respectively. Hence, the overall response rate was of 60.1% and can, therefore, be considered ‘good’.

Table 3.2 - Questionnaires' response rate

Year	Group 2	Group 3	Total
Students who consented taking part in the study	358	294	**652**
Questionnaires	186	206	**392**
Response rate	52.0%	70.1%	**60.1%**

3. *Focus groups*

Focus groups are particular regarding purpose, size, composition and procedures. They are used to gather opinions and their purpose is to enable the researcher to understand ‘why’ participants think or feel the way they do (Bryman & Bell 2011). These groups can be composed of 4 to 12 people who have something in common and the interviewer, who plays a facilitator role, encourages all kinds of comments regarding a focused discussion (Krueger 2015). Focus groups are recommended when the researcher is looking for a range of opinions about some issue, practice, or pilot-test. It is not the goal to achieve consensus and it is even desired that ideas are varied and emerge from the group. Also, they are considered helpful when quantitative data has been collected, and researchers want to understand what some of the results are attributed to (Krueger 2015).

In the case of this research, focus groups were used to gather additional qualitative data that helped the researcher to understand better the opinions and feelings of participants regarding: being anonymous or identifiable, the experiences of using a likert or dichotomous scale, and – most importantly – the effects of having their personal networks online. That is, referring back to the conceptual model and its link to the data collection (Figure 3.4), focus groups helped to get insightful discussions with the users of PeerWise regarding the ratings they gave, and how these were affected by different user profiles and rating scales. Further, focus groups provided invaluable information that not only helped to respond to the ‘hows’ and ‘whys’ but also guided the researcher in analysing specific online interactions that otherwise would have been overlooked. Therefore, the analysis of the quantitative part of the experiment partly guided the topics to be discussed in the focus groups. Yet, these discussions also made the researcher return to and analyse different aspects of the data.

As described above, focus groups were conducted when the semester finished to avoid low attendance rates due to final exams, but before students obtained their final grades for the module in order to minimise the impact of grades on their opinions. However, unavoidably, students had already received the marks concerning their PeerWise assignment, which may have had an impact on the discussions. Further, the method for recruiting participants was that, at the end of the questionnaire, students were asked if they would be willing to attend a focus group to further discuss their experience with PeerWise. Therefore, to respect their privacy, only those pupils who answered affirmatively were later contacted with the details of dates, venues, and incentives. Regarding incentives, all students were offered – and given – a £10 Amazon voucher if they participated. Vouchers were approved by the Ethics Committee of the Management School. Appendix 3.3 shows the script of the focus group.

The ideal number of participants for a focus group is between five and eight. Also, the literature suggests to plan for three or four focus groups and only schedule more if *saturation* (i.e. "the point where the researcher has heard a wide range of ideas and is not getting new information") is not reached (Krueger 2015, p.23). In this research, the plan was to have – per group – four focus groups of five people each, scheduling up to seven students per session, in case some dropped out. However, the reality was slightly different but still favourable, as there were six focus groups in total, three in Group 2 and three in Group 3 (see Table 3.1).

3.3.4 Limitations

As outlined in Section 3.1, in line with the post-positivist standpoint adopted, it is required that the researcher should reflect on the limitations of the methods of data collection. These have been highlighted throughout the chapter, but it was considered relevant to further explain the most significant ones.

Firstly, quasi-experiments have the advantage of gaining authenticity, but scientists can sometimes struggle to prove that there was a change and that this was caused by a determined event (Campbell & Ross 1968). Thereafter, when a change is detected, it is legitimate to ascribe it to the quasi-experimental treatment "provided consideration is given to plausible rival explanations of the differences, with supplementary analyses being added to eliminate these where possible" (Campbell & Ross 1968, p.37). In order to tackle this limitation, the researcher made use of three sources of data to be able to triangulate findings. Further, as one of the data sources was focus groups, ultimately participants were able to reflect on the 'hows' and 'whys' of their experiences.

Secondly, as outlined in Table 3.1, Group 1 only contained one source of data (i.e. online interactions), as opposed to Groups 2 and 3, which also had questionnaires and focus groups. Unfortunately, this condition could not be mitigated as Group 1 was not 'controlled' by the researcher and access to the students was not possible due to ethical restrictions. Consequently, it was not possible to triangulate the findings from the first cohort, and assumptions had to be made relying only on the data that was available.

A final limitation is that, in order to make information comparable, it used the same environment to imitate features of both SNSs and CCs. Therefore, although the findings of this research are considered to be applicable for a wide range of online platforms where UGC is shared, it should not be forgotten that the experiment is set in an online educational environment and generalisations outside this must be made with caution.

3.3.5 Ethical considerations

Before conducting the research, the quasi-experimental design on PeerWise, the questions from the survey and the discussion topics for the focus groups, were all scrutinised and later approved by the Ethics Committee of the Management School. The following ethical considerations should be highlighted:

Signing-in students with their real identities (i.e. Name_LastName)

By default, PeerWise is designed so that students choose their own usernames, although all interactions are made anonymously. Hence, the data retrieved from Group 1 of the quasi-experiment showed that more than half of students created a username that made them identifiable, either by the presence of their first or last names. At the point of signing-in, students had no previous knowledge about PeerWise, so they did not know that the software would anonymise all interactions. Therefore, this meant that more than half of the group did not mind being identifiable by their peers. Moreover, PeerWise is set up so lecturers can identify students[21]. Hence, PeerWise would be catalogued as a nonymous environment, as users are always linked to their 'official' identifiers (Zhao et al. 2008). Therefore, given that more than half of the students from Group 1 did not mind being identifiable and because students from the three cohorts were always aware that the module staff could identify them, signing everyone with their real identities and making most interactions visible was not considered a significant change. Still, all students from Groups 2 and 3 were given the option to contact the module leader or the researcher if they wanted to be given an anonymous ID. However, for both cohorts, no one made this request.

PeerWise scores do not affect the module's assessment

It should be once again highlighted that nothing from within PeerWise affected the students' grades. That is, the following scores were *not* taken into account for the assignment: number of correct or incorrect answers, receiving high or low ratings, reputation scores, number of awarded badges, or appearing on the leaderboards. Therefore, even if making students identifiable affected the ratings they gave to each other, this in no way affected their grades. As previously mentioned, PeerWise was only used in the selected module as a supporting tool for students to have online discussions; it was not the main source for learning and the interactions that took place there did not affect the grades in any way.

Rating and commenting is optional

PeerWise is designed so that rating and commenting are optional, and the experiment did *not* modify these rules. Likewise, neither ratings nor comments were required for the assignment and therefore had no effects on student grades. This point is important because, by being identifiable, some students might find stressful to rate or comment on their classmates' questions; especially if some of these were authored by their friends. For this reason, PeerWise remained unchanged on these features. Therefore, students

[21] When setting up a module within PeerWise, lecturers/administrators need to upload a list of the students' identifiers (e.g. emails or student IDs), to prevent strangers from accessing the group.

could choose to solely answer questions, and this would remain anonymous on the website. Nevertheless, if they chose to rate or comment, their names would accompany their ratings and posts.

The only personal question in the questionnaire was optional

To obtain users' personal networks, the questionnaire required students to name three of their peers who they knew from class, and then say how frequently they met outside the lecture. To respect student privacy, the question was made in two parts: indicating three people they knew was made mandatory (for those who wished to take part in the survey), while saying how frequently they saw their peers was optional. This caused some data to be incomplete, but the privacy of students came first.

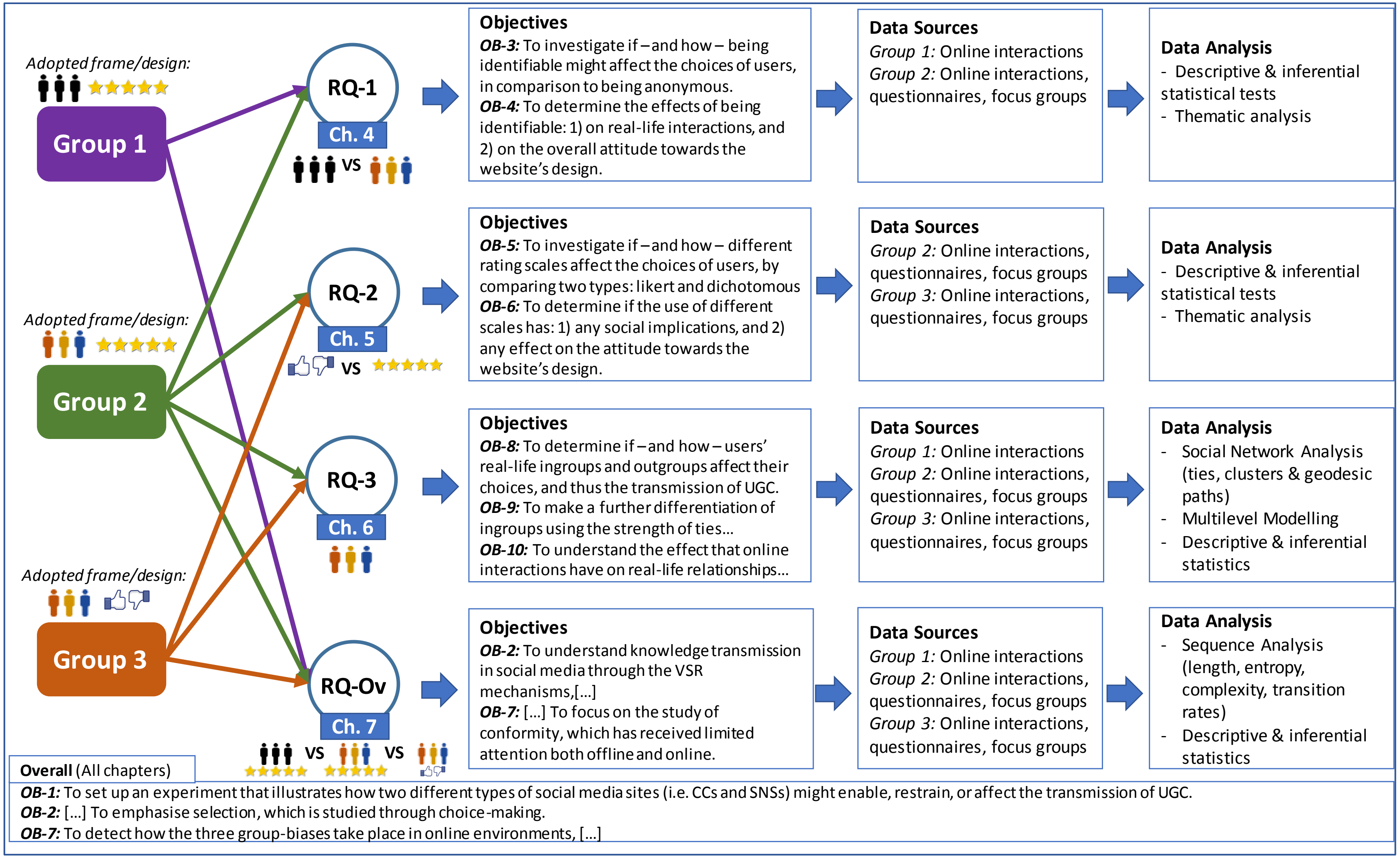
Participants could drop-out at any time during the study

Everyone who participated in the research was given an information sheet and was required to fill-in a consent form, for every step of data collection: extracting online data, responding to the questionnaire, and participating in the focus groups. Further, no one was ever contacted without having expressed interest in the following step of data collection. Finally, being mindful of participants' self-consciousness, they were clearly advised that they could withdraw from the research at any time (Conway & Lance 2010), even after having taken part.

3.4 PROCEDURE FOR DATA ANALYSIS

Figure 3.5 below presents a summary of the sources of data collection linked to the methods for data analysis, separated by research question. As can be observed, the quantitative and qualitative data coming from the three quasi-experimental conditions were analysed with the use of four methods. Hence, the following sub-sections describe each of these four methods: 1) statistical tests, 2) thematic analysis, 3) social network analysis, and 4) sequence analysis. These methods are broadly introduced here, with each chapter of results having a separate methods section in which the particular variables and tests performed are described in detail.

Figure 3.5 – Quasi-experimental conditions, sources of data collection and methods for data analysis, per research question & chapter



3.4.1 Statistical tests

Statistical tests were conducted in SPSS on the quantitative data coming from the online interactions and close-ended questions from the survey. Regarding online interactions, these were first used to obtain descriptive statistics for every group, in order to compare them. For instance, the first step towards addressing both RQ-1 and RQ-2 (Chapters 4 and 5) is a comparative table showing the average number of questions, answers, ratings and comments per group, followed by some indicators per capita (i.e. per participant). Further, a histogram was used to contrast the average ratings between groups, and this is accompanied by the overall mean, standard deviation, and both a t-test (for parametric data) and a Mann-Whitney test (for non-parametric data) (Field 2013). T, after comparing the ratings for all the authored questions, additional between-group comparisons were made only for the identical questions across groups (see Figure 3.3). Finally, to test all the variables within the three groups of biases (content, prestige, and conformity), t-tests, ANOVA, and regressions were performed, depending on whether variables were categorical or continuous (Field 2013).

In addition, the quantitative data from the questionnaires were automatically coded in Qualtrics and later imported and analysed using both Excel and SPSS. As mentioned, questionnaires aimed to understand the attitudes that students had towards being anonymous or identifiable, utilising a likert or dichotomous scale, and the presence of their personal networks on the website. Thus, the effect that these attitudes had on the overall experience of the website was measured using Spearman’s rho correlation coefficient, which measures the strength of association between two ranked variables (e.g. Brendgen et al. 2005). Lastly, the correlations were accompanied by histograms and other basic statistics.

3.4.2 Thematic Analysis

Thematic analysis was used to make sense of the qualitative data that came from the open-ended questions of the questionnaire and the focus groups. It should be noted that, although the thematic analysis was only used for Chapters 4 and 5, triangulation was used in all chapters, and the quotes from students were used throughout the analysis to support, question, and explain the ‘why’ of findings.

Before conducting thematic analysis, the qualitative data from the questionnaires and focus groups had to be combined, in order to be coded and arranged by themes. The data coming from the questionnaires was already typed and could be easily extracted from Qualtrics. However, information collected from the focus groups had to be transcribed. According to the methodological literature, there are a number of ways of managing the data obtained in the focus groups, such as transcripts, abbreviated transcripts, notes, and memory (Krueger 2015). This research made use of *abridged transcripts*, which consist of transcribing only the relevant parts of conversations. In other words, the following were not transcribed: introduction, first question, excessive moderator directions, and comments that do not directly relate to the purpose of the study (Krueger 2015). These types of transcripts are meant to be conducted by someone who is familiar with the study and, in this case, the person who designed the experiment and conducted the focus groups was in charge of performing the transcriptions (i.e. the researcher).

After transcribing the focus groups, there are a wide range of techniques that can be used to make sense of the data. For instance, content analysis could be used to quantify content within predetermined categories, or narrative analysis could be applied to analyse data that is sensitive to the temporality in which participants relate their experiences (Bryman & Bell 2011). However, if the goal is to identify, analyse and report patterns in the data, as in this research, then thematic analysis is the most suitable tool (Braun & Clarke 2008). This type of analysis aims to observe and cluster key themes that occur during the discussion with participants, in order to generate a coherent interpretation of the collected data. Further, thematic analysis "can be applied across a range of theoretical and epistemological approaches" (Braun & Clarke, 2006, p. 78), and is thought of being "the most useful in capturing the complexities of meaning within a textual data set" (Guest et al. 2012, p.10). For all these reasons, thematic analysis was considered the most appropriate tool for dealing with the qualitative data of the quasi-experiment.

Moreover, thematic analysis can be conducted in a theory-driven manner, where themes are developed based on the available knowledge about the issue that is being studied. Conversely, it can also be done in a data-driven approach, when there is little information available on the topic; or it can be done in a hybrid manner, combining both theory and data-driven (Fereday & Muir-Cochrane 2006). The data-driven approach is considered to be more useful when "investigating an under-researched area", or when "working with participants whose views on the topic are not known" (Braun & Clarke 2008, p.83). Thus, given that most of the issues investigated in this research are under-studied (e.g. website designs and conformity), this research made greater use of the data-driven approach.

The analysis began with three main categories, each one corresponding to one of the research questions of this thesis: 1) anonymous versus identifiable users, 2) likert versus dichotomous rating scales, and 3) personal networks. Afterwards, second-level themes emerged from the comments of participants. For this, the research followed the six phases for conducting thematic analysis suggested by Braun and Clarke (2008): familiarising with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report. The software used to conduct thematic analysis were *Excel* and *NVivo*.

It should again be highlighted that qualitative data not only served for triangulation and thematic analysis. Actually, thanks to the insights from participants, it was possible to examine specific aspects of the online interactions and social network analysis that would have otherwise been overlooked. Thus, there was a recurrent dialogue between the quantitative and qualitative parts of the thesis, in which both helped to obtain a better understanding of the other.

3.4.3 Social Network Analysis

Social Network Analysis (SNA) is a research field that concerns the analysis of networks of individuals or organisations (Thelwall 2009). It focuses on the investigation of relationships between social entities, and on the patterns and effects of these relationships (Wasserman & Faust 1994). SNA was used to investigate the third research question, which is addressed in Chapter 6.

Networks have been defined as "a set of actors connected by a set of ties" (Borgatti & Foster 2003, p.992). For this research, the actors or *nodes***,** are online users. Further, ties connect pairs of actors and can be *directed* (e.g. Twitter followings, which do not need to be reciprocal) or *undirected* (e.g. Facebook friends, where the friendship needs to be mutual), and *dichotomous* (e.g. present or absent, as with Twitter follower) or *valued* (e.g. measured regarding the strength of friendship, as with Facebook's smart lists) (Borgatti

& Foster 2003). In the case of this research, the data that were collected through the survey involved directed, valued connections. However, for the analysis, ties were assumed to be undirected; valued connections. That is, if a student A declared that she saw student B frequently, it was assumed that B would also see A frequently, unless specified differently by student B.

To obtain undirected, valued ties, the survey questions that dealt with networks and the strength of ties were extracted from Qualtrics in a CSV[22] format, and then converted from the original matrices to nodes using *MATLAB*. Once the undirected ties were obtained, the analysis was as follows. First, the rating-averages of friends (i.e. undirected, valued ties) and non-friends (i.e. absent ties) were compared. Second, the rating-averages among different valued ties were contrasted (e.g. Granovetter 1973; 1983). That is, rating-averages were compared between absent (non-friends), weak (acquaintances), intermediate (friends), and strong ties (close friends).

Third, as an exploratory analysis, clusters were created using the software *Gephi*. The goal of cluster analysis is to group multivariate data and, unlike classification, it is used when there is no *a priori* group information (Dean 2016). Namely, data is meant to be unsorted and the aim is to investigate if there are any groups in the data and, if so, how many and what they look like. This analysis was undertaken to explore the effect of relationships and clusters in ratings, and provided very insightful findings that can be explored in future research with the use of multilevel modelling (e.g. Tranmer et al. 2014).

Fourth, also in an exploratory manner, undirected ties were used to obtain geodesic distances. These can be defined as the shortest path between any two nodes (i.e. students) within a network (Jackson 2008). Geodesic distances present the advantage of being always consistent because, unlike clusters, there is always the same distance between two people and it always corresponds to the shortest path. Moreover, as opposed to clusters where a person is allocated into only one group, with geodesic distances the same person can have as many connections as the number of members in the network. That is, people are not allocated to one or the other cluster, but instead have a definite geodesic distance with every other member of the network. Given that geodesic distances provided more

[22] CSV: comma-separated values.

‘reliable’ results, these were used to compare the declared friendships from the survey with PeerWise’s ‘followers’ (see Appendix 3.1).

3.4.4 Sequence Analysis

Sequence Analysis (SA) is used to detect patterns in categorical sequences, focusing on the state of the sequence, where “the position of each successive state receives a meaningful interpretation” (Gabadinho, Ritschard, Mueller, et al. 2011). SA was used to investigate conformity to the whole network and as a way of answering the overall research question, which is addressed in Chapter 7.

SA was used to compare the levels of conformity among the three groups. For this analysis, only the ratings per question were used and were analysed for each group with the *TraMineR* package, in the *R* software (Gabadinho, Ritschard, Studer, et al. 2011). With this tool, some indicators such as within entropy, complexity index, and transition rates were used, concepts that are explained in detail in Chapter 7. Nevertheless, it is relevant to mention that SA was applied in this thesis for the first time to the study of social media and ratings, and the findings obtained with this method have proven to be insightful in understanding and comparing the effects that different website designs have on the ratings of online users.

Conclusion

This chapter has presented the philosophical assumptions and research methods adopted in this thesis. The research is rooted in a post-positivist standpoint which, much like positivism, is characterised by objectivity, causality and generalisation. However, it acknowledges that individuals cannot achieve objectivity in a perfect manner. In line with this view, a quasi-experiment was designed and a mixed- and multiple-methods approach was used to gather quantitative and qualitative data. An educational online platform, PeerWise, was utilised to conduct the quasi-experiment, in which changes to the website’s design were made regarding user profiles and the rating scale. Moreover, data were collected from three main sources: retrieved online interactions from PeerWise, questionnaires, and focus groups. Four broad methods for data analysis were used: statistical, thematic, social network, and sequence analyses. Finally, it should be highlighted that all research activities in this thesis were approved by the Ethics Committee of the University of Sheffield's Management School.

The next four chapters provide results that target the posed research questions. Chapter 4 examines how two different user profile set-ups affect the ratings (i.e. choices) of users through the use of statistical and thematic analyses. Chapter 5 investigates the effect that two different rating scales have on the choices of participants, also via statistical and thematic analyses. Chapter 6 explores the effect of the presence of personal networks in online environments through comparing conformity towards the whole and personal networks, and different strengths of ties making use of social network analysis, clustering and geodesic distances. Lastly, Chapter 7 makes an overall comparison of the three quasi-experimental conditions presented in Groups 1, 2 and 3, by performing sequence analysis to uncover rating patterns among the three cohorts.

CHAPTER 4: THE EFFECT OF DIFFERENT LEVELS OF SELF-PRESENTATION ON CHOICES

As explained in Chapter 2, when an individual is in the presence of others, they will try to obtain information about her, such as her socio-economic status, background, trustworthiness, and competences. Hence, when an individual is before others, she will have several reasons for seeking to control the impression she gives (Goffman 1959). In online environments, identity can be seen as the extent to which users reveal themselves or the amount of personal information that sites allow to be shared (Kietzmann et al. 2012). In particular, self-presentation is a component of identity by which an individual tries to make an impression on others (Goffman 1959).

The internet gives individuals the opportunity to alter their identity to an extent that would not be possible in face-to-face communication by allowing them to change aspects of what Goffman referred to as 'personal front', which comprises their age, gender, and appearance (Suler 2002). Therefore, by allowing features like disembodiment and anonymity, online environments allow for a new means of identity construction (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008). More concretely, within social media, individuals are free to design their physical forms (e.g. avatars, human, animal, hybrids), gender, and any wished symbolic connotations (Schau & Gilly 2003).

Regarding the two types of social media under consideration – CCs and SNSs – the first is characterised by having a low level of self-presentation, while the second has a high level (Kaplan & Haenlein 2010). Regarding user profiles, which are arguably responsible for the different levels of self-presentation, the most distinctive characteristic that differentiates CCs from SNSs is that the former allows anonymity, mainly because the emphasis is on the content being shared. Thus, in CCs users are usually not required to create a profile if they only wish to browse information, although some sites do require very basic profiles when people want to post comments (Kaplan & Haenlein 2010). Conversely, in SNSs users are defined as 'nonymous' or non-anonymous (i.e. identifiable) because in these sites relationships are anchored to offline environments, which means that users are meant to be known by others to a certain extent (Zhao et al.

2008). Hence, SNSs require users to create a profile, encourage the creation of 'friends' lists', and allow individuals to navigate their list of connections to see other people's information (Amichai-Hamburger & Vinitzky 2010).

This chapter aims to determine if the characteristic levels of self-presentation of CCs and SNSs affect the transmission of information. The different levels of self-presentation are studied through the comparison of anonymous and non-anonymous (i.e. identifiable) user profiles, and the impact on information transmission is seen through the choices of users (i.e. their ratings).

4.1 FIRST RESEARCH QUESTION AND OBJECTIVES

This chapter addresses the first research question of the thesis: *How are the choices of users affected by different levels of self-presentation occurring from diverse user profiles?*

To address this research question, the chapter focuses on tackling objectives 3, 4, and 7, outlined below. It should be noted that the numbering of objectives is kept consistent with Chapters 2 and 3 (Table 2.4 and Figure 3.5, respectively). Further, Figure 4.1 shows how these objectives fit with the conceptual model of the thesis.

- **OB-3:** To investigate if – and how – being identifiable might affect the choices of users, in comparison to being anonymous.
- **OB-7:** To detect how the group-biases[23] take place in online environments […]
- **OB-4:** To determine the effects of being identifiable: 1) on real-life interactions, and 2) on the overall attitude towards the website's design.

[23] Only content- and prestige-based biases are examined in this chapter. Conformity is not analysed given that the personal networks of participants were not collected for Group 1, and thus the between-group comparison cannot be performed. Hence, objective 7 is partly fulfilled in this chapter, and is further studied in Chapters 6 and 7.

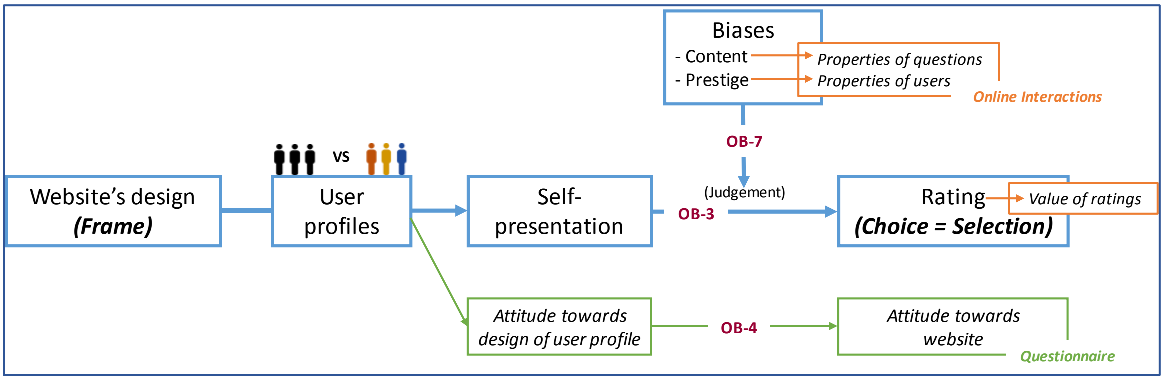


Figure 4.1 - Objectives of first research question and link with the conceptual model

4.2 METHOD

4.2.1 Participants

As explained in Chapter 3, all participants were final-year undergraduate students. The total population used in this chapter is of 664 students: 367 females and 297 males. Group 1 – where users interacted in PeerWise anonymously – comprised 295 participants, with 267 based in the UK and 28 in China. Users in Group 2 – who were signed in with their real identities – comprised 369 students, 330 based in the UK and 39 in China.

4.2.2 Quasi-experimental set-up and data collection

To fulfil the objectives outlined, two quasi-experimental conditions were designed in PeerWise. Then, the online interactions generated on both groups were compared (e.g. questions, answers, ratings, comments, replies, badges, leaderboards). In Group 1, students were allowed to choose their usernames, but the website was set up (per default) so that all online interactions were anonymous. In contrast, in Group 2, students were signed in with their real identities (i.e. '*Name_LastName')* and all their online interactions were identifiable. Figure 4.2 shows the information that was displayed, on each group, when accessing the question of a peer, in order to answer, rate, or comment on it. As can be seen, being identifiable meant that the names of the 'author' of the question and the 'raters' were public. It should be noted that this image only presents the impact that being anonymous/identifiable had on ratings, but these conditions affected other elements and interactions within the website. Appendix 3.1 contains several screenshots that show all the changes that took place on PeerWise.

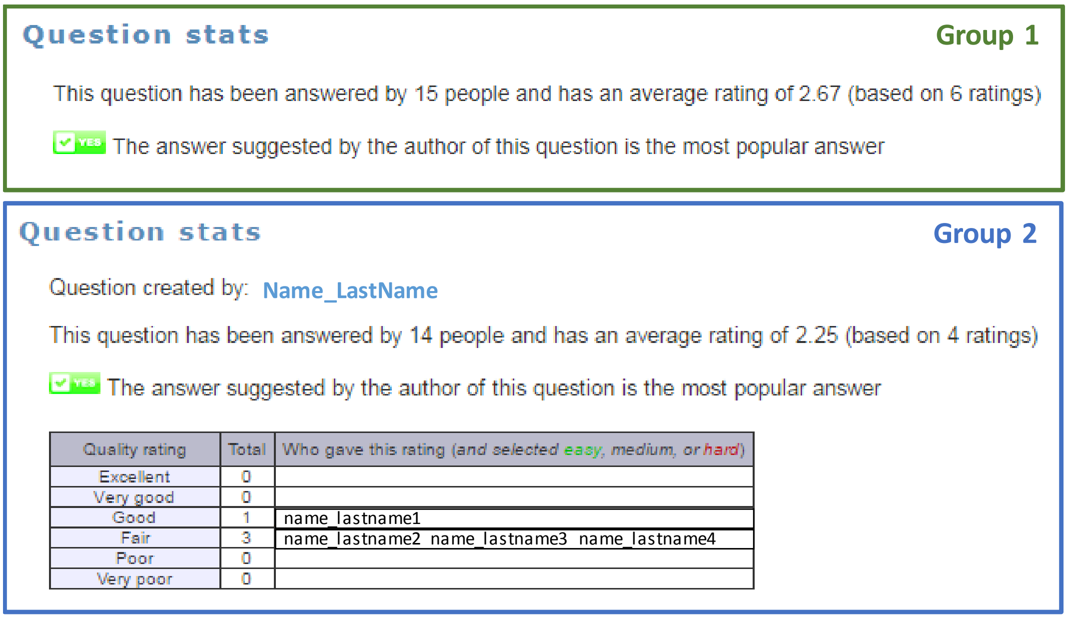


Figure 4.2 – Tackling a question on PeerWise when anonymous (Group 1) and when identifiable (Group 2)

For both groups, online interactions were extracted and compared using SPSS. In addition, some of the participants from Group 2 also took part in a questionnaire (N=186) and three focus group (N=17). The quantitative data from the questionnaire were analysed using Excel and SPSS, and the qualitative data coming from the open-ended questions of the survey and the focus groups were analysed with Excel and NVivo.

4.2.3 Procedure

The procedure for data analysis and description of the variables was as follows:

A. *Investigating the effects of two user profiles on ratings (Figure 4.1, OB-3)*

First, to have an overview of both groups, the primary online interactions from both cohorts were contrasted with the use of average, maximum, and per capita values. Second, a between-group comparison of the ratings was made through the use of histograms, and then through an *independent sample t-test*, which was used when the participants are split into two groups corresponding to two different experimental conditions (Field 2013). It should be noted that there is an ongoing debate on whether ordinal rating scales (i.e. likert scales) can be considered normal distributions, and if it is correct to perform parametric tests on them (de Winter & Dodou 2010). Although it is common in the social sciences to treat likert scales as normally distributed, statisticians suggest to start by performing a normality test. However, as a rule of thumb, normality tests such as *Kolmogorov-Smirnov* should only be used for 'medium-size' datasets, as they tend to give false positives for smaller datasets and false negatives for larger ones (Field 2013). For this reason, given that the present data set was considered large, it was decided to check normality through the visual inspection of the distributions and their *Quantile-Quantile (Q-Q) plots*. Nonetheless, to increase validity, the researcher also performed the *Mann-Whitney-Wilcoxon* non-parametric test to determine if the difference in means was significant (de Winter & Dodou 2010).

Third, after conducting the between-group comparison for all ratings, this analysis was repeated for the ratings given to the 15 identical questions in Groups 1 and 2 (see Figure 3.3). Fourth, the findings from the statistical tests were triangulated with the thematic analysis conducted on the comments from the participants of the second cohort. As mentioned in Chapter 3, the creation of themes followed the six phases for conducting thematic analysis suggested by Braun and Clarke (2008).

B. *Exploring which group of biases has the strongest effect on ratings (Figure 4.1, OB-7)*

As mentioned, for this chapter only content- and prestige-based biases are examined. Conformity is not analysed given that the personal networks of participants were not collected for Group 1, and thus the between-group comparison is not possible. Furthermore, as can be seen in Figure 4.1, content-based biases comprised variables that captured specific properties of questions, whereas prestige-based biases concerned variables reflecting specific properties of participants. Although Chapter 3 briefly described the three groups of independent variables, a list[24] with each of their descriptions is presented below. In addition, the type of each variable is highlighted: continuous, categorical, and ordinal. Further, categorical and ordinal values can sometimes be binary, when a characteristic is either present or absent.

Content-based biases → Specific properties of questions

- *Question's date (Continuous):* Date when the question was posted online. Extracted from PeerWise.

[24] This list shows the variables studied in this section. However, Chapter 4 completes the set.

- *Question's time (Continuous):* Time when the question was posted online. Extracted from PeerWise.
- *Question's link (Categorical-Binary):* Whether the question contained a web link (URL) or not (e.g. to a TED Talk or a YouTube video). Obtained for every question using MATLAB.
- *Question's reference (Categorical-Binary):* Whether the question included a reference or not (e.g. "Smith (2004) argues, …). Obtained for every question using MATLAB.
- *Question's number of words (Continuous):* Number of words in the question. References and links were deducted from the total word count. Calculated for every question using MATLAB[25].
- *Question's number of sentences (Continuous):* Number of sentences comprising the question. References and links were deducted from the total count of sentences. Calculated for every question using MATLAB.
- *Question's number of characters (Continuous):* Total number of characters (letters and numbers) in the question. The characters contained in references and links were deducted from the total count. Calculated for every question using MATLAB.
- *Question's ARI readability index (Continuous):* The Automated Readability Index (ARI) is a readability test designed to evaluate how readable a text is (Smith & Senter 1967). The formula to obtain this index is: *4.71 * (characters / words) + 0.5 * (words / sentences) – 21.13*. This index was calculated for every question using Excel.
- *Question's CLI readability index (Continuous):* The Coleman-Liau Readability Index (CLI) assesses how comprehensible a script is (Colmer 2013). The formula to obtain it is as follows: $0.0588*L - 0.296*S - 15.8$, where L is the average number of letters for 100 words and S the average number of sentences per 100 words. This index was calculated for every question using Excel.
- *Question's tag (Categorical-Binary):* Whether the question contained a tag or not. PeerWise allowed students to tag their questions; that is, to categorise them by topic, case study, or company (e.g. 'Environmental issues' or 'Coca-Cola'). The

[25] It should be noted that the only software used in the thesis that the researcher did not use herself was MATLAB. This tool was used with the help of Dr Umberto Esposito, from the Computer Science department of the University of Sheffield.

tag was extracted from PeerWise and was later converted into a binary variable using Excel.

- *Question's explanation (Categorical-Binary):* Whether the question included an explanation or not. When students authored a question, they had the option of adding an explanation. This was shown after their peers had attempted the question, but before they could rate or comment on it. In most cases, the explanation gave a rationale for the 'correct' answer as determined by the author (see Appendix 3.1 for all screenshots of PeerWise). The explanation was extracted from PeerWise and was later converted into a binary variable using Excel.
- *Question's number of alternatives (Ordinal):* All questions were multiple-choice, and students had the option to give others between two and five alternative answers to choose from. The number of alternatives was extracted from PeerWise.
- *Rating's date (Continuous):* Date-stamp of each rating given to a particular question. Extracted from PeerWise.
- *Rating's time (Continuous):* Time-stamp of each rating given to a particular question. Extracted from PeerWise.

<u>*Prestige-based biases → Specific properties of participants*</u>

<u>*Online-gained prestige*</u>

- *Author's distinct badges (Continuous):* PeerWise gave students a number of distinct and repeated badges. The maximum number of distinct badges that a student could have was 25. Extracted from PeerWise.
- *Author's repeated badges (Continuous):* There was no limit for the number of repeated badges a student could get. The maximum number obtained in Group 1 was 369, and 363 in Group 2. Both types of badges were linked to users' 'online reputation score', which was one of the options by which students could select questions in Group 2, where student identities were made visible. Extracted from PeerWise.

<u>*Real-life prestige*</u>

- *Author is admin (Categorical-Binary):* As explained, at the start of each semester the Admin staff of the module posted some questions, identical throughout the cohorts (see Figure 3.3). All groups were aware that the module's staff authored these initial questions – even in the first cohort when all online interactions were anonymous – as students were made aware of this during the lecture.

- *Author's nationality (Categorical):* Given that there over 50 different nationalities but three quarters of the students were comprised in only two of them, these categories were formed: UK, China, Others. This data were retrieved from the University's database.
- *Author's gender (Categorical-Binary):* Male or female. Data retrieved from the University's database.

Regarding the procedure, given the mix of continuous, ordinal, categorical, and binary variables, it was difficult to determine which variable – or which group of biases – was having the most significant effect on ratings, as there is no 'right' way to do this. On the one hand, a 'common' method is to use hierarchical or stepwise methods to include or delete variables in a model and observe the change in the coefficient of determination (Field 2013; Stone et al. 2016). On the other hand, other scholars suggest it is best to compare the standardised regression coefficients (Thompson 2009; Bhalla 2015; Stone et al. 2016). To increase reliability in the comparison of variables, the following procedure was followed: first, each of the previously outlined independent variables was assessed individually against the dependent variable (i.e. ratings). Ordinal and categorical values were examined with a t-test, when binary; and with an ANOVA when they presented 3 or more states. Continuous variables were assessed with the correlation coefficient. A bootstrap of 1,000 samples was performed for each of the tests, to make them more robust. These analyses were performed for both groups and were later compared in a table. Second, the variables that proved to have the highest individual effect on the means were included in a regression. Given that many variables were highly correlated (e.g. words with characters, sentences, and readability indexes), only those with the strongest effect on ratings were considered for the regression. Third, regressions were performed for each group, using the most significant variables of each type. Regressions were executed using the *backwards stepwise method*, so variables were initially 'forced' into the model and then deleted one by one if they were highly correlated or not significant (Field 2013). Fourth, standardised coefficients were used to rank the predictor variables of each group. Fifth and finally, the ranking was compared among groups.

C. *Determining the effects of the attitude towards being identifiable on the attitude towards the website (Figure 4.1, OB-4)*

In the questionnaires, students were asked about their perception towards being identifiable and towards the whole website, by evaluating both items with a likert scale. Hence, in order to test the effect of students' attitude towards being identifiable on their views about the website itself, two analyses were conducted. Firstly, Spearman's rho was obtained, and then an ordinal regression was performed. Spearman's rho measures the strength of association between two ranked variables. The given coefficient varies between -1 and 1, and the closest it is to these values, the stronger the correlation. Conversely, the closer it is to zero, the weaker the correlation (Lund & Lund 2015).

4.3 RESULTS

The results are organised around the chapter objectives. First, the focus is on comparing the ratings of Groups 1 and 2, and thus determining whether the different levels of self-presentation obtained through different user profiles produced any impact. This objective is achieved first through the statistical tests outlined above, and then through triangulating the findings with the qualitative data. Second, the variables comprised within content- and prestige-based biases are ranked regarding their effect on ratings, to explore which evidenced the strongest group-bias. Third, the effect of attitudes towards identifiability on perceptions of the website is tested, and this analysis is also supported by the comments of participants.

4.3.1 Comparing the impact that different levels of self-presentations have on choices

Summary of online interactions of Group 1 and Group 2

Although the focus of this research is on ratings (i.e. choices), it is essential to have an overview of the primary online interactions for both groups, given that these were also affected by the quasi-experimental conditions. Table 4.1 shows the number of questions, answers, ratings, comments and replies, among others.

Table 4.1 – Summary and comparison of online interactions between Groups 1 and 2

	Group 1	**Group 2**
Students (active in PeerWise) **Based in the UK** **Based in China**	Total: 295 UK: 267 CH: 28	Total: 369 UK: 330 CH: 39
Total no. of questions (non-deleted)	2,104	2,525
Max. no. of questions per student	29	19
Ave. no. of questions per student	7.10	6.78
Total no. of answers	34,018	34,379
Max. no. of answers per student	683	644
Ave. no. of answers per student	120.39	96.04
Total no. of ratings	25,927	25,976
Ave. no. of ratings per student	87.89	70.40
Mean rating of all questions	2.64	2.94
Total no. of comments	12,414	13,805
Max. no. of comments per student	398	485
Ave. no. of comments per student	44.20	38.37
Answered but not rated	8,074 = 23.7%	8,403 = 24.4%
Answered but not commented	21,603 = 63.5%	20,574 = 59.8%
Rated but not commented	13,529 = 52.2%	12,171 = 46.9%
Max. no. of total badges per student	369	363
Ave. no. of total badges per student	59.11	48.81
Max. no. of distinct badges per student	25	25
Ave. no. of distinct badges per student	18.06	16.75

As can be seen, although Group 2 had 74 additional students and over 400 more authored questions than Group 1, it had almost the same number of answers and ratings. That is, there were fewer answers and ratings per capita in the second cohort. Still, the proportion of students that answered a question and subsequently rated it was almost the same for both years at about three quarters. Further, the proportion of students who answered a question and then left a comment was higher in Group 2. Therefore, these differences could suggest that, by being identifiable, students became more selective about which questions to answer, or perhaps more self-conscious. In addition, those who were anonymous seemed to have a preference to rate questions after answering them, whereas identifiable users appeared to prefer leaving a comment. Finally, in Group 2 the average rating of all questions was higher than in Group 1, which suggests that anonymous pupils were 'tougher' when evaluating the questions of their peers. The following sub-sections focus on the comparison of ratings.

Between-group comparison of all ratings

Figure 4.3 shows the visual comparison of all ratings of Groups 1 and 2. Before conducting such analysis, the researcher performed a Kolmogorov-Smirnov normality test for both datasets and, as expected, it turned out to be significant in both cases, given the large amount of data they contain. Moreover, by making a visual inspection of both of the histograms shown below and the Q-Q plots, it was concluded that the data followed a normal distribution (see Appendix 4.1 for the normality tests of all ratings). Hence, an independent sample t-test was conducted. Group 1 had a mean rating of M=2.64 (SD=1.17, N=25,927), while G2 had an average rating of M=2.94 (SD=0.89, N=25,976). The mean difference, of Δ=0.30, was very significant at $t(48{,}410) = -32.88$, $p<.001$. Further, to increase validity, an independent-samples Mann-Whitney non-parametric test was also performed, and this also concluded that the mean difference was highly significant: $U = 398{,}147{,}787$, $z = -37.81$, $p < .001$.

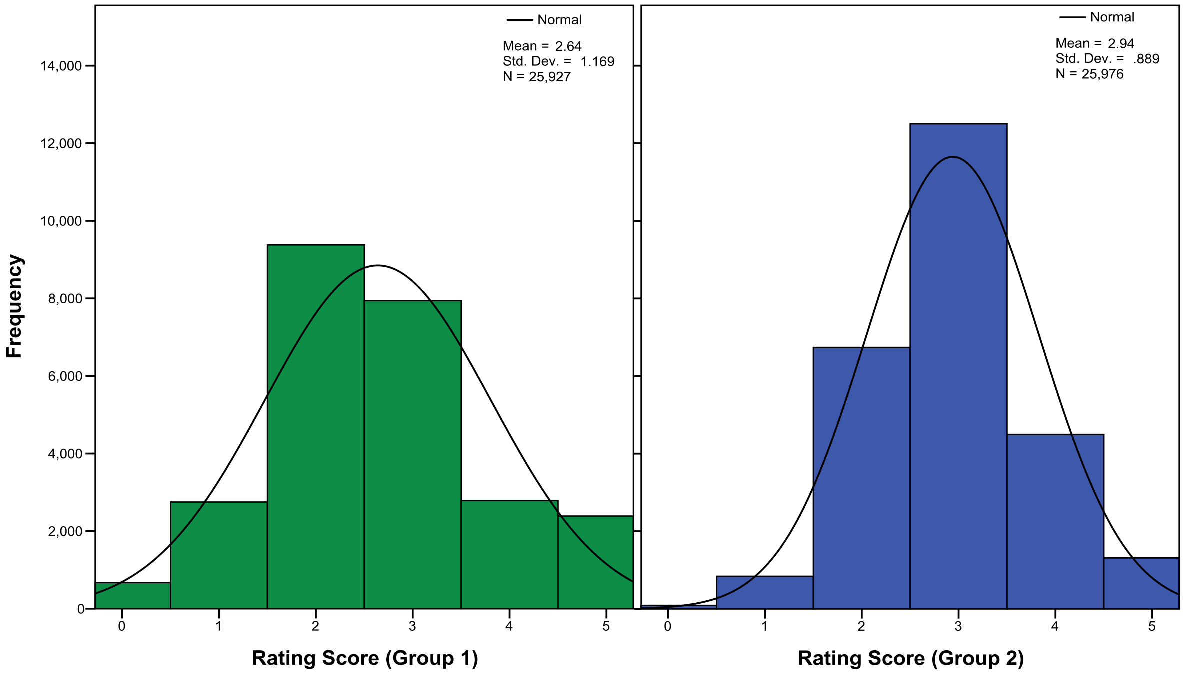


Figure 4.3 - Visual comparison of all ratings in Groups 1 and 2

As described, both the parametric t-test and the non-parametric Mann-Whitney test provided evidence to support the claim that the difference in rating means between

Groups 1 and 2 was strongly significant. However, what is even more remarkable is the comparison between the two distributions shown in the histograms presented above. As can be seen, students from Group 2 not only had a higher average rating than those of Group 1, but their manner of rating was very different. That is, it was not only a change in the central average but in the range of values. The most prominent finding, perhaps, is that rating scores of zero, one, and five decreased substantially for Group 2. To examine this claim further, the table below summarises the frequency and percentages for each of the values within the rating scale. It can be seen that Group 1 had a more even distribution and its modal score was 2 (Fair), with 36.2% of the users giving this rating. On the contrary, ratings from Group 2 gathered around the middle values, and the modal rating among this cohort was 3 (Good), with 48.1% of the total ratings.

Table 4.2 - Rating distribution breakdown, for Groups 1 and 2

	Group 1			**Group 2**		
	Frequency	**Percent**	**Cum. Percent**	**Frequency**	**Percent**	**Cum. Percent**
0 (Very poor)	673	2.6%	2.6%	89	0.3%	0.3%
1 (Poor)	2,750	10.6%	13.2%	839	3.2%	3.6%
2 (Fair)	9,380	36.2%	49.4%	6,738	25.9%	29.5%
3 (Good)	7,945	30.6%	80.0%	12,502	48.1%	77.6%
4 (Very Good)	2,791	10.8%	90.8%	4,496	17.3%	94.9%
5 (Excellent)	2,388	9.2%	100.0%	1,312	5.1%	100.0%
Total	**25,927**	**100.0%**		**25,976**	**100.0%**	

Table 4.2 shows that when users interacted anonymously in PeerWise, they used the whole spectrum of the rating scale. In contrast, when online interactions were associated with users' real identities, they avoided extreme values; especially those with a negative connotation. Moreover, the most frequently used value shifted from 2 (Fair) in Group 1, to 3 (Good) in Group 2, indicating that students in the second cohort were rating in a 'nicer' or less critical way. This last point further confirms the results of the parametric and non-parametric tests. Therefore, the statistic tests, histograms, and the rating distributions all indicate that there is a significant difference in the way people rate, which is 'harsher' when anonymous and 'nicer' when identifiable.

Between-group comparison of 15 identical questions

Similar to the overall findings, when considering the common questions, the average rating of Group 2 (M=2.93) was found to be higher than that of Group 1 (M=2.91). However, due to the small mean difference (Δ=.014) and the comparatively smaller samples (N_1=443 and N_2=485), the mean difference of these 15 questions was not significant, at t(898) = -.237, p = .813.

However, when the histograms of both years are visually inspected (see below), they show that distribution for Group 2 is again more centralised, with 'Good' (3) being used almost twice as frequently used as ratings of 2 and 4. In contrast, Group 1's histogram shows a more even distribution and a more use of frequency for extreme values. This is reflected in a higher standard deviation for the first cohort (SD=.922) compared with the second set of students (SD=.845). Thus, although the difference in means of the 15 repeated questions is not statistically significant, they show the same rating patterns as the comparison for all questions, which was highly significant.

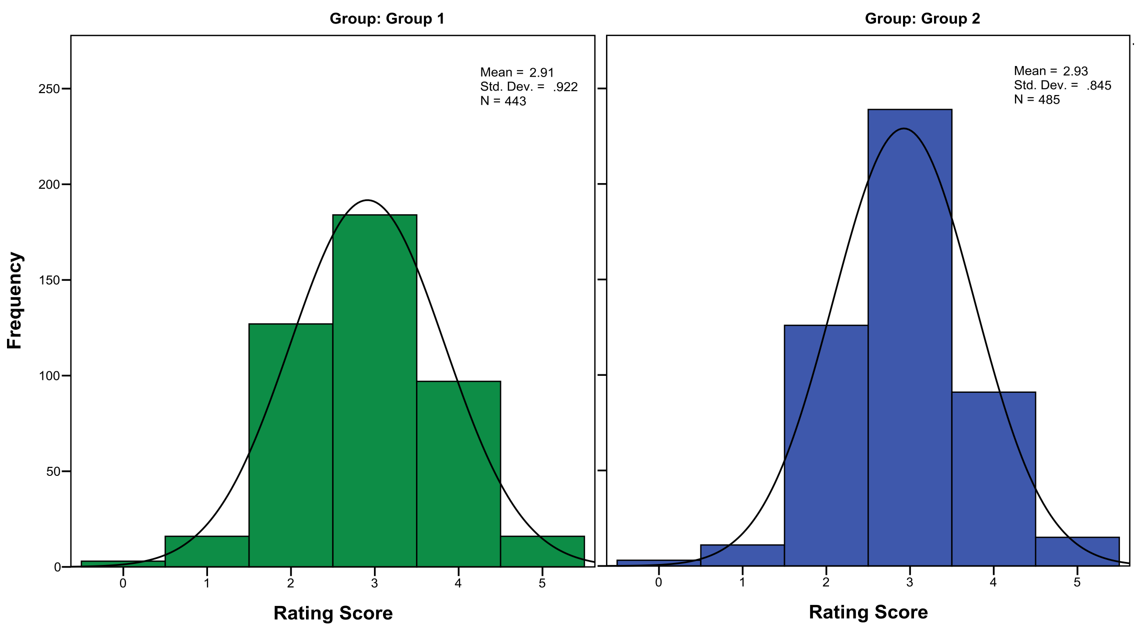


Figure 4.4 - Histograms of 15 identical questions of Groups 1 and 2

Triangulation of findings – interacting online when being identifiable

When triangulating the quantitative findings above with the thematic analysis of participants' qualitative data, it was found that several students felt that being identifiable had an impact on the way they rated. Specifically, some of the reasons why 'negative' ratings were avoided can be seen in Table 4.3, which presents one representative quote from each theme:

Table 4.3 - Thematic analysis regarding the avoidance of negative ratings when identifiable

Theme	Original quotes from participants *[sic.]*
Fear of revenge	"*I am aware of situations where somebody had rated a question not very highly (their fair opinion) and the question author retaliated by searching all of the rater's questions, rating all of them at the bottom rating - this isn't helpful or fair*" [sic.] (Quest-G2, enri651).
Collaboration and reciprocity	"*Peer lwise seemed to be Almost a popularity contest where users would use their own friends to gain reputation, score and to get ahead. The only way to get ahead was to do the same. Those who didn't collaborate were at a disadvantage*" [sic.] (Quest-G2, rdph008)
Repercussions in real-life	"*Though you want to answer your friends questions for whatever reasons, it promotes bias and people don't give real answers. Even if they are not a close friend, you won't speak overly negative of another classmate as they may take it personally and can see who said so*" [sic.] (Quest-G2, elra972).
Felt obliged to rate positively	"*Made people biased in rating marks and also when giving feedback. If you knew the person you didnt want to give them negative feedback even if you didnt like the question. In addition, you wouldnt except bad feedback off people you knew. I had heard many cases where people were even mentioning this in lectures*" [sic.] (Quest-G2, imwh215)
Peer-pressure	"*Because it would be more fair to answer the questions anonymously. As i mentioned before, they rated my questions as excellent and they told me to rate their questions as excellent as well but i didnt want to do that. So if it was anonymous, Peerwise would be 100% better*" [sic.] (Quest-2, nist698).
Would rather not rate	"*I wouldn't rate, rather than rate a question badly… But I think I would have done the same if I were anonymous*" [sic.] (FG2-G2, lala535).

The table above explains some of the 'whys' regarding the significant difference in the mean ratings of Groups 1 and 2. Further, it also suggests why this difference made Group

2 seem ‘nicer’. Specifically, it is worth highlighting that many of the themes partly address Objective 4 (i.e. ‘determine if being identifiable affected real-life interactions). As can be seen, five of the six themes above indicate that the rating had to do with a social context, to some extent. Hence, regardless of whether it was a negative impact (e.g. fear, repercussions or pressure) or a positive incentive (e.g. collaboration and reciprocity), it seemed that being and having peers identifiable affected the manner in which users rated.

Some users were aware of the tendency for ratings in Group 2 to conform towards the centre (prevailing the mode of 3): *“I feel like the ratings didn't hold any real significance, most questions only achieved a 2.5 average rating”* (Quest-G2, asad101). However, when asked, participants had different ideas on how they would overcome this situation. On the one hand, some just suggested to anonymise it: *“I would keep the scale as it is, however I would suggest anonymising the process to avoid people just using their friends to comment on and rate their questions”* (Quest-G2, camu282). On the other hand, it was also recommended that they could remain identifiable and merely change the rating scale. Here again, some students believed a broader scale would be the solution *“I think the options 'fair' and 'good' ended up being chosen more frequently than they should have. By giving a sliding scale of 1-10 you'd receive a more accurate rating”* (Quest-G2, ieje546). Others however, thought fewer options might be more suitable: *“People converge to the middle ratings rather than giving the full range. Also, as students don't actually know what is good/correct for the assessment, rating something as poor or excellent is not necessarily representative. Like/dislike is easier to just see that person's opinion on the question”* [sic.] (Quest-G2, enri651).

4.3.2 Impact of content- and prestige-based biases on choices

This section focuses on understanding the possible impact that biases could have had on the ratings of anonymous and identifiable users. Further, it aims at ranking all variables included in the biases to determine which group had the strongest impact within each group. Once again, it should be noted that, even though this thesis studies three groups of biases (content, prestige, and conformity), only the first two have been considered in this chapter because Group 1 lacks the questionnaire which collected personal networks.

Individual Statistical tests

Table 4.4 presents a summary of the individual tests performed on each independent variable in relation to the dependent variable: ratings. As explained in Section 4.2.3, this was the first step before deciding which variables to include in the regression that is later used to compare the standardised coefficients, so as to rank the variables.

Table 4.4 - Summary of individual tests for variables of Groups 1 and 2

Independent variable	Var. type // Test	Test results	
		Group 1	**Group 2**
Question's date	Continuous // R^2	R^2 = 0.1% of variation in rating score. Significant at F(1)=27.43, p < .001.	R^2 = 1.7% of variation in rating score. Significant at F(1)=439.52, p < .001.
Question's time	Continuous // R^2	R^2 = 0.0% of variation in rating score. Not significant at F(1)=1.47, p=.225	R^2 = 0.1% of variation in rating score. Significant at F(1)= 21.70, p<.001
Question's number of words	Continuous // R^2	R^2 = 2.7% of variation in rating score Significant at F(1) = 710.74, p < .001. // *Plot shows that questions with 300+ words never got ratings of 0 or 1.*	R^2 = 3.1% of variation in rating score. Significant at F(1) = 820.53, p < .001. *Plot shows that questions with 200+ words never received a 0, while 300+ words only got 3, 4, or 5.*
Question's number of sentences	Continuous // R^2	R^2 = 2.5% of variation in rating score. Significant at F(1) = 673.64, p < .001. //*Questions with 25+ sentences never got scores of 0 (very poor) or 1 (poor).*	R^2 = 2.6% of variation in rating score. Significant at F(1) = 706.65, p < .001. *Plot shows that questions with 15+ sentences never got ratings of 0.*
Question's number of characters	Continuous // R^2	R^2 = 2.7% of variation in rating score. Significant at F(1) = 718.66, p < .001. // *Questions with 1,500+ characters did not get ratings of 0 or 1.*	R^2 = 3.4% of variation in rating score. Significant at F(1) = 911.04, p < .001. *1,000+ characters did not get ratings of 0; 1,500+ characters only received 3, 4, or 5.*
Question's ARI read. index	Continuous // R^2	R^2 = 0.1% of variation in rating score. Significant at F(1) = 32.46, p < .001.	R^2 = 0.3% of variation in rating score. Significant at F(1) = 77.64, p < .001.
Question's CLI read. index	Continuous // R^2	R^2 = 0.2% of variation in rating score. Significant at F(1) = 42.61, p < .001.	R^2 = 0.5% of variation in rating score. Significant at F(1) = 123.80, p < .001.
Question's link	Categorical (binary) // t-test	Link absent: M=2.63, SE=.031 Link present: M=2.83, SE=.007 Mean difference: Δ= -.203 Significant, t(25,925)= -6.65, p<.001.	Link absent: M=2.85, SE=.008 Link present: M=3.05, SE=.008 Mean difference: Δ= -.201 Significant, t(25,540)= -18.33, p<.001.
Question's reference	Categorical (binary) // t-test	Reference absent: M=2.60, SE=.008 Reference present: M=2.91, SE=.020 Mean difference: Δ= -.314 Significant, t(25,925)= -15.01, p<.001.	Reference absent: M=2.91, SE=.006 Reference present: M=3.12, SE=.014 Mean difference: Δ= -.212 Significant, t(25,974)= -14.13, p<.001.

Question's tag	Categorical (binary) // t-test	Tag absent: M=2.58, SE=.026 Tag present: M=2.65, SE.008 Mean difference: Δ= -.063 Significant, t(2,631)= -2.31, p<.05	Tag absent: M=2.87, SE=.023 Tag present: M=2.94, SE=.006 Mean difference: Δ= -.072 Significant, t(1,827)= -3.10, p<.005
Question's explanation	Categorical (binary) // t-test	Explanation absent: M=2.04, SE=.188 Explanation present: M=2.64, SE=.007 Mean difference: Δ= -.599 Significant, t(69.21)= -3.181, p<.005.	Explanation absent: M=2.64, SE=.097 Explanation present: M=2.94, SE=.006 Mean difference: Δ= -.300 Significant, t(112.72)= -3.12 , p<0.005
Question's number of alternatives	Ordinal // ANOVA	2 alternatives: M=2.23, SE=.129 3 alternatives: M=2.54, SE=.026 4 alternatives: M=2.61, SE=.011 5 alternatives: M=2.68, SE=.010 Significant at F(3) = 17.39, p<.001.	2 alternatives: M=2.94, SE=.011 3 alternatives: M=2.93, SE=.012 4 alternatives: M=2.91, SE=.010 5 alternatives: M=2.99, SE=.012 Significant at F(3) = 9.50, p<.001.
Rating's date	Continuous // R^2	R^2 = 0.2% of variation in rating score Significant at F(1)= 53.31, p < .001.	R^2 = 2.4% of variation in rating score. Significant at F(1)= 649.26, p < .001. *No question got a rating of 5 after the assignment was over (late November)*
Rating's time	Continuous // R^2	R^2 = 0.1% of variation in rating score. Significant at F(1)= 26.88, p < .001.	R^2 = 0.0% of variation in rating score. Not significant
Author is (student / admin)	Categorical (binary) // t-test	Author student: M=2.64, SE=.007 Author Admin: M=2.89, SE=.043 Mean difference: Δ= -.250 Significant, t(461.8) = -5.73, p<.001.	Author student: M=2.94, SE=.006 Author Admin: M=2.93, SE=.038 Mean difference: Δ= -.014 Not significant at p= .725
Author's distinct badges	Continuous // R^2	R^2 = 1.6% of variation in rating score. Significant at F(1)= 417.61, p < .001.	R^2 = 0.5% of variation in rating score. Significant at F(1)= 119.52, p < .001.
Author's repeated badges	Continuous // R^2	R^2 = 0.6% of variation in rating score. Significant at F(1)= 168.67, p < .001.	R^2 = 0.1% of variation in rating score. Significant at F(1)= 16.40, p < .001.

Some of the findings from the table above were also present in the comments of users. To mention a few, in one of the focus groups (FG2-G2) all participants agreed that the date and time when they had posted their question had an impact on the ratings they received. Moreover, some individuals also reflected on content-related elements: *"A couple of the questions I authored included short videos. These seemed to get higher ratings than the questions involving articles. I don't know whether these questions were of higher quality or whether respondents liked the fact that there was visual content and the questions took less time to answer"* [sic.] (Quest-G2, lala535). Moreover, other users also mentioned badges and reputation scores influenced ratings: *"…very few people would give a high rate if the author doesn't have a high reputation…"* [sic.] (Quest-G2, yuto987).

However, although analysing individual variables separately and reading student comments provided many insights, to determine which group of biases have the strongest impact on ratings a regression was performed, and standardised coefficients were compared. Then, as outlined, when variables explain the same issue, only the ones with the most substantial effect on ratings are included in the regression. For instance, in Table 4.4, by comparing words, characters, sentences, and indexes, the biggest variation in the mean was due to the number of characters. Hence, only this variable was used for the regression, whereas number of words, sentences, and indexes were discarded. In the same manner, from all the variables regarding dates and times, rating's date had the strongest relationship with ratings. Finally, between distinct and repeated badges, the first one had greater impact in both cohorts.

Using standardised coefficients to rank the effect of the variables on ratings

In order to make an equal comparison, a linear regression model was estimated for each group. Both models can be found in Appendix 4.2. As a summary, in Group 1, 5.7% of the variation was explained by the linear relationship between ratings and the predicting variables, and this was highly significant overall ($F(8, 25{,}918) = 194.44$, $p < .001$). Likewise, in Group 2, 7.1% of the variation was explained by the linear relationship between ratings and the predicting variables, and this was again highly significant ($F(8, 25{,}967) = 249.76.03$, $p < .001$). It should be noted that regression models are only used to determine which groups of biases are stronger for each group. It is not the intention to obtain a model that predicts ratings, nor to determine if this model should be linear or logarithmic. Namely, the researcher is not claiming that the relationship between variables is linear, nor that these models have sufficient explanatory power to make predictions.

Table 4.5 presents a summary of the ranking of the standardised coefficients, for both groups. These have been converted to absolute values to measure their effect on ratings, regardless of whether the effect is positive or negative. Moreover, they have been sorted in a within-group ranking, to determine which was the strongest for each group. It should be noted that, although most variables showed the same effects in the individual tests and in the regression, in a few cases their significance changed due to the association with other variables. This issue is known as *confounding*, which is a situation in which the relationship between a variable and the outcome is distorted by the presence of another variable (Lengerich 2016). For instance, ‘question’s tag’ had been significant (at $p < .005$) for Group 2 when considered alone, yet appeared not to be significant in the regression and was therefore removed for both groups. Likewise, questions authored by ‘Admin’ had not been individually significant for the second group but it turned out that it had a certain degree of significance in the regression, and was hence included.

Table 4.5 - Ranking of variables using standardised coefficients (Groups 1 and 2)

	Group 1		Group 2	
Rank	**Independent variable**	**Stand. Coeff. (|Abs|)**	**Independent variable**	**Stand. Coeff. (|Abs|)**
1	Author's Distinct Badges	.147	Rating’s date (by week)	.150
2	Qs’ number of characters	.146	Qs’ number of characters	.149
3	Author’s Admin	.100	Author’s Distinct Badges	.106
4	Qs’ Reference	.075	Author’s Admin	.083
5	Rating’s date (by week)	.040	Qs’ Reference	.072
6	Qs’ Link	.033	Qs’ number of alternatives	.065
7	Qs’ number of alternatives	.020	Qs’ Link	.052
8	Qs’ Explanation	.016	Qs’ Explanation	.021

From the table above, it can be seen that for Group 1 the aspect having the strongest impact on the ratings of anonymous users was ‘distinct badges’ (a prestige-based bias), followed by the question’s number of characters (content-related), and then questions authored by the Admin (prestige). Conversely, for the cohort where students were identifiable, the variable that had the strongest effect on ratings was the rating’s date, followed by the question’s number of characters (both content-related biases), followed by the number of an author’s distinct badges (prestige-based). Therefore, these results suggest that Group 1 complied with what is expected by research performed in both group biases and social media: people perceive information from those that have some degree of prestige as being more relevant or of better quality. Yet Group 2 does not seem to confirm these results. However, as can be read in the comments from students, by being identifiable, different kinds of dynamics arose; some of which involved friendships (i.e. conformity-based biases). These dynamics are studied in the following chapter.

Highlighting differences between Groups 1 and 2

Based on the statistical tests performed on each variable (Table 4.4) and the comparison of the ranking of variables within each group (Table 4.5), it is possible to comment on some of the most significant differences between the two cohorts studied in this chapter. It should be noted that some of these findings were 'emergent' and can therefore be slightly unrelated to the topics covered in the literature review. Still, all results are shown at this point as they respond to the outlined objective *"OB-3: To investigate if users would rate differently when their identity is known to others than when it is not"*.

1) *The ratings per week presented distinct trends*

In both cases, the average rating before the assignment was due was higher than after the assignment. This finding suggests that, although students were told that ratings would not count towards their grades, they still rated differently over the weeks when the use of PeerWise was mandatory. Nonetheless, Group 1's ratings went down as the assignment was due, while those of Group 2 went up (see Figure 4.1). A possible explanation for this can be found in one of the themes from Table 4.3, 'collaboration and reciprocity' where, as a student mentioned, those who did not collaborate were at a disadvantage. Moreover, the way in which participants described the 'tactics' used to rate seemed to put them in one of two categories. On the one hand, some users described the group working 'collaboratively': *"Because a lot of cases occurred where classmates would collaborate with each other about what was the correct answer to questions etc…"* (Quest-G2, rdph008). On the other hand, other participants referred to what could be described as a 'competitive strategy': *"I felt some people rated other questions poorly, so that they would get higher ratings"* (Quest-G2, cysm998). Therefore, a plausible cause for ratings going in opposite directions as the deadline of the assignment approached, is that anonymous users tend to make use of a 'competitive strategy', whereas identifiable users might have found more success with 'collaboration'.

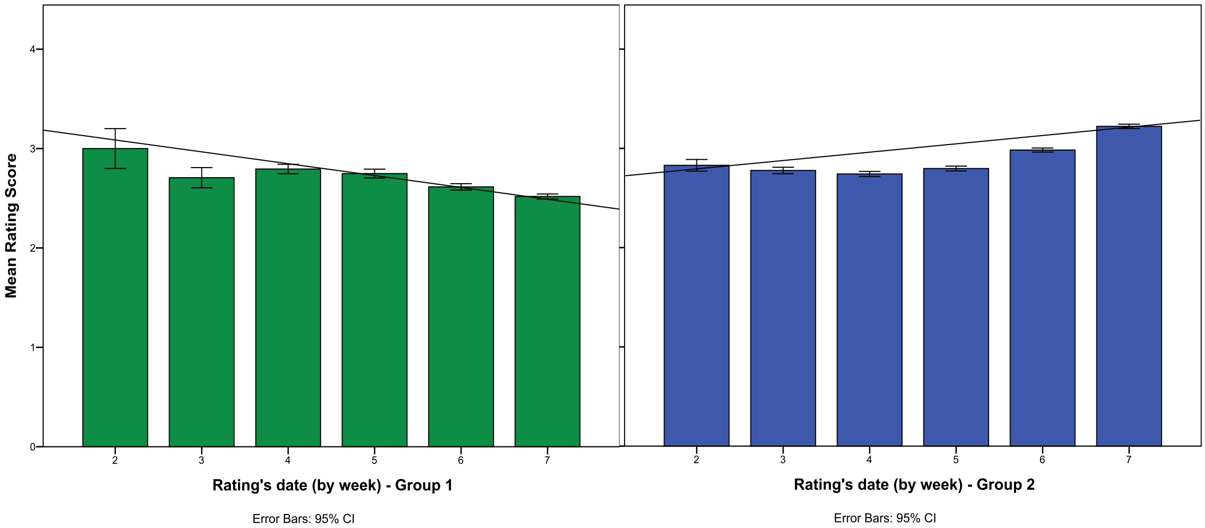


Figure 4.5 - Ratings by week during mandatory use of PeerWise, Groups 1 and 2

2) *Admin questions were rated differently*

As explained, even participants who interacted anonymously knew that the module's staff had posted the first questions. Further, in all cohorts, students were aware that the admin staff would be monitoring the module and could identify all individuals. However, an interesting finding was that students who were anonymous rated the questions authored by the admin above the group average, while identifiable students rated them below their group's average (see Table 4.4). In a way, it seemed as if students behaved in a 'nicer' way towards the module staff when only the staff members knew their identities. Conversely, students behaved slightly 'harsher' with staff members when their ratings were also identifiable by their classmates. This finding is somehow contradicting with the overall ratings because, as has been said, overall Group 1 seemed to rate 'harsher' than Group 2. Moreover, as will be discussed later, this finding is also in contradiction with other studies in exploring anonymity. Further, although the comments of students offered no insights as to why this might have happened, one of the researcher's supervisors offered the following plausible explanation: *'No one wants to be seen as the teacher's pet'*. Hence, as will be suggested later, future research could explore the role of admin-pupil interactions in the context of absence/presence of classmates, or other observers.

3) Reliance on online-gained prestige

As seen in Table 4.5, author's distinct badges was the predictor that played the most prominent role for Group 1, whereas it was only the third most important predictor in Group 2. This result suggests that, when users are anonymous, they rely heavily on what has been defined in this thesis as 'online-gained prestige' (e.g. badges, reputation scores, leaderboards, rankings). Conversely, when users were identifiable, content-based biases have a greater role, although the quotes from participants seem to suggest that friendships are important too. These will be explored in the following chapters of results.

Including two additional prestige-based variables

As outlined in Chapter 3 and the method section of this chapter, there was a recurrent dialogue between the quantitative and the qualitative data. That is, the findings of one led to further analysis in the other, and so on. Hence, although the analysis of biases was meant only to include the variables outlined in Table 4.4 and Table 4.5 but, prompted by some of the student comments, it was decided that two more variables should be included. The qualitative input from participants made the researcher analyse a previous phase to the rating of questions, which was not considered, namely the selection of questions. In a way, this slightly deviates from the focus of this study, which is on ratings. However, selecting questions is still part of making a choice, and it directly affects ratings as students could only rate questions they had previously chosen and answered. Therefore, if the selection of questions suffered any changes due to students being identifiable, it is also worth studying.

The first comment came from the questionnaire, where a participant mentioned that, by being identifiable, students selected questions based on the author's identity: *"…it is very easy to dismiss what could have been a stimulating question just by seeing the author's name"* (Quest-G2, lyco891). Moreover, a second comment came from a British student in one of the focus groups, who avoided questions if she believed the authors were international students, based on the expectation that their level of English would be poor: "*Also, this is going to sound really bad, but if I see that a Chinese person has posted a question, and I know their English isn't good, I'm not going to be inclined to answer their question if I don't understand what they're asking me*" (FG3-G2, ieje546). Finally, at the end of that focus group, when the researcher asked students if they would have preferred to be anonymous or identifiable and why, another student added: "*Yeah, and [while using real names] some people might avoid Chinese's questions, just by seeing their names they would assume their questions would not be good enough. And [being anonymous] would just take away all of that*" (FG3-G2, lyco891).

Based on the above comments, it was decided first to explore the percentage of targeted questions between nationalities. Table 4.6 shows a 3x3 matrix of each group in which the nationalities of both the author of the question and the student who answered the question are shown. It should be noted that these questions exclude the answers given to the questions authored by the module staff. As can be seen, when identifiable, students from China targeted more questions from UK students, while the reverse did not happen. On the contrary, the percentage of answers by British students to questions with Chinese authors dropped from 27.3% in Group 1 to 20.2% in Group 2, when users were identifiable. This represented 1,677 fewer answers from British to Chinese students from Group 1 to Group 2.

Table 4.6 - Comparative matrix of nationalities of answered questions, Groups 1 and 2

Author from →	**CHI**	**Other**	**UK**	**% of answers**
Group 1 – Answer from:				***N = 33,403 answers***
CHI	49.4%	16.5%	34.0%	**32.7%**
Other	27.5%	26.9%	45.6%	**19.4%**
UK	27.3%	17.0%	55.7%	**47.9%**
Total Answers	***34.6%***	***18.8%***	***46.6%***	***100.0%***
Group 2 – Answer from:				***N = 33,784 answers***
CHI	46.2%	16.2%	37.7%	**36.0%**
Other	24.6%	29.5%	45.9%	**24.7%**
UK	20.2%	19.8%	59.9%	**39.3%**
Total Answers	***30.6%***	***20.9%***	***48.4%***	***100.0%***

Furthermore, given that when users were identifiable the choice of questions selected was affected by personal factors, it was decided to include two more variables and study their effect on ratings. The two variables, nationality and gender, were catalogued as 'real-life prestige'. In theory, if they happened to be significant, it should only be for identifiable users. Table 4.7 presents the individual tests for both variables:

Table 4.7 - Introducing nationality and gender variables to Groups 1 and 2

Independent variable	Var. type // Test	Test results	
Prestige-based bias		**Group 1**	**Group 2**
Author's gender	Categorical (binary) // t-test	Males: M=2.63, SE=.011 Females: M=2.64, SE=.010 Mean difference: Δ= -.012 Not significant at p=.422	Males: M=3.03 SE=.009 Females: M=2.87, SE=.007 Mean difference: Δ= .161 Significant, t(25,489) = 14.32, p<.001
Author's nationality [UK, China, Others]	Categorical // ANOVA	UK: M=2.66, SE=.010 China: M=2.68, SE=.013 Others*: M=2.49, SE=.017 Significant, F(2)=47.99, p<.001 **30 different nationalities*	UK: M=2.97, SE=.008 China: M=2.95, SE=.010 Others*: M=2.87, SE=.012 Significant, F(2)=24.24, p<.001 **39 different nationalities*

As expected, gender was not significant for anonymous students, whereas it had a small but very significant impact on students who were identifiable. As can be seen, in Group 2 men got significantly higher ratings than women, on average. Conversely, nationality was significant for both groups, which is somehow unexpected as in Group 1 users were not identifiable. Still, ratings changed from one year to the other, as in Group 1 Chinese students got the highest mean ratings, whereas in the second cohort British students obtained the highest ratings, on average. These findings, together with the ones from Table 4.6, evidence that personal aspects influenced students' ratings when these were identifiable. Particularly, British and male students appear to have benefitted the most in terms of ratings when being identifiable.

Given the findings above, a regression model was run for each cohort. This was followed by a ranking of standardised coefficients in order to determine where the two new variables fitted in relation to the others. Regarding the models, Group 1 produced a significant regression model overall (F(9, 25,480) = 172.614, p < .001) with an R^2 of 5.7%. Similarly, Group 2 also had a significant regression (F(9, 25,481) = 237.85, p < .001) with an R^2 of 7.7%. It should be noted that adding the two new variables, gender and nationality, the coefficient of determination remained the same for Group 1, but improved by 0.6% for Group 2. Thus, this gives further proof of gender and nationality having an impact when users interact and are identifiable, as opposed to anonymous.

The ranking of all variables is displayed in Table 4.8. As can be seen, results are very similar to the ones shown from the previous models, as the top-5 variables stayed the same for both groups. Moreover, as found with the individual tests, gender was only significant for Group 2, while nationality was significant for both cohorts, and with almost identical effects on ratings.

Table 4.8 - Ranking of variables through standardised coefficients, including gender and nationality, Groups 1 and 2

	Group 1		Group 2	
Rank	**Independent variable**	**Std. Coeff. (|Abs|)**	**Independent variable**	**Std. Coeff. (|Abs|)**
1	Author's distinct badges	.147	Rating's date (by week)	.150
2	Qs' number of characters	.146	Qs' number of characters	.149
3	Author's Admin	.100	Author's distinct badges	.106
4	Qs' Reference	.075	Author's Admin	.083
5	Rating's date (by week)	.040	Qs' Reference	.072
6	Author's nationality	.038	Qs' number of alternatives	.065
7	Qs' Link	.033	Author's gender	.059
8	Qs' number of alternatives	.020	Qs' Link	.052
9	Qs Explanation	.016	Author's nationality	.037
10	Author's gender	*Not signif., p=.423*	Qs' Explanation	.021

Some possible explanation of why nationality had a small, but significant, impact on the ratings from Group 1 – when it should have been not significant, given that participants were anonymous – are as follows. First, it could be because, as declared from some students during the focus groups, they seemed to have used WhatsApp groups to get support from the classmates they knew. *"To be honest, whenever I would write a question, I would just message my friends and be like 'can you answer it?' I mean, why wouldn't I? That's just clever"* (FG3-G2, ieje546). Alternatively, as some students also mentioned, it was common for classmates (who had the same location and, possibly, the same nationality) to work towards the assignment in groups: *"I saw people in the library one day, and they were just answering each other's questions"* (FG1-G2, eaio573). Moreover, PeerWise generates an individual ID for each authored question, which appears in the hyperlink. Therefore, students could have just copied and sent the link among their personal networks, even when being anonymous. Further, it may be that users employed the 'follow' function to subscribe to the questions of known peers. However, as focus groups were not conducted for Group 1, there is no certain way to know this.

To sum up, it has been so far shown which biases were the strongest for each group, and how they differed between groups. Prestige-based biases predominated in Group 1, while content-based ones played a more prominent role in Group 2; although conformity-based biases have yet to be tested (see the following chapter).

4.3.3 Attitude to being identifiable and its impact on the website's perception

This last section focuses on the attitudes and comments that students shared about being identifiable. As has been mentioned, Group 1 was subject to the limitation that it did not have a questionnaire or focus groups. Therefore, the findings from this section are only from students in Group 2. This section first covers the overall attitude that students had on regarding being identifiable, followed by the overall perception they had towards PeerWise. Both quantitative and qualitative data are used. The final part measures the effect that the attitude towards being identifiable had on views about the website, using only quantitative data and with the use of the Spearman's rho correlation coefficient.

Attitude towards being identifiable

Firstly, during the questionnaire, students were asked if being able to identify one another with usernames *'Name_LastName'* was helpful or not. If very unhelpful takes the value of one and very helpful takes the value of five, the mean of Group 2 was of 3.29. Then, taking into account this average and the breakdown of all values (see Table 4.9 below), more than half of the group thought being identifiable was helpful. Still, it should also be noted that all categories got almost the same percentages:

Table 4.9 - Attitude towards being identifiable, Group 2

Being identifiable	No. of students	Percentage
Very unhelpful	27	14.6%
Unhelpful	26	14.1%
Neutral	40	21.6%
Helpful	50	27.0%
Very helpful	42	22.7%
Total	***185***	***100.0%***

Moreover, students were asked, if they had a choice, whether they would have preferred their classmates and themselves to be signed in with their real names or, instead, using anonymous IDs. From the 185 students that responded these two questions, 37.8% replied they would like to be able to identify others, while 62.2% said they would not like to know their classmates' identities. Similarly, 35.1% said they would like to continue to be identified by others by being signed in with their real names, while 64.9% said they would rather not be identified by their peers. Therefore, only about a third of students would have kept their identifiable usernames, whereas the other two-thirds would instead use an anonymous ID. Further, although it was not a significant difference, it is worth highlighting that five students changed their answer when the question regarded the identity of others, as opposed as to themselves. Hence, this showed that a few participants liked the idea of knowing their peer identities within the website, but would have kept their own interactions invisible from others.

One of the students who preferred being identifiable suggested that:

> *"by showing my own first and last name I think that it provides personable benefits to the replying and discussion process. Having myself anonymous on peerwise; I would feel disconnected from others and more sensitive to negative criticism. Having names on peerwise avoids this and I feel more comfortable knowing the person I am in discussion with. Therefore, negative discussion could be a risk if names aren't shown"* (Quest-G2, ergo113)

Conversely, one comment that summarises the 65% that would rather be anonymous focused on some of the specific matters that are the focus of this thesis, clearly demonstrating that at least some students are aware of these issues:

"To avoid bias and ratings based on the relationship with the author. Anonymous would be ideal, as students then tend to rate based on the quality of the question instead" (Quest-G2, emha308).

Overall attitude towards the website

Towards the end of the questionnaire, participants were asked how helpful they thought PeerWise was. To enable comparison, Figure 4.6 presents a histogram of their answers, compared with those concerning being identifiable. As can be seen, these histograms differ significantly. Although barely half (49.70%) of the students thought being identifiable was helpful, the majority (86.2%) thought this about PeerWise. Conversely, 28.7% of students thought being identifiable was not helpful, compared to only 4.2% of students perceiving this about the website itself.

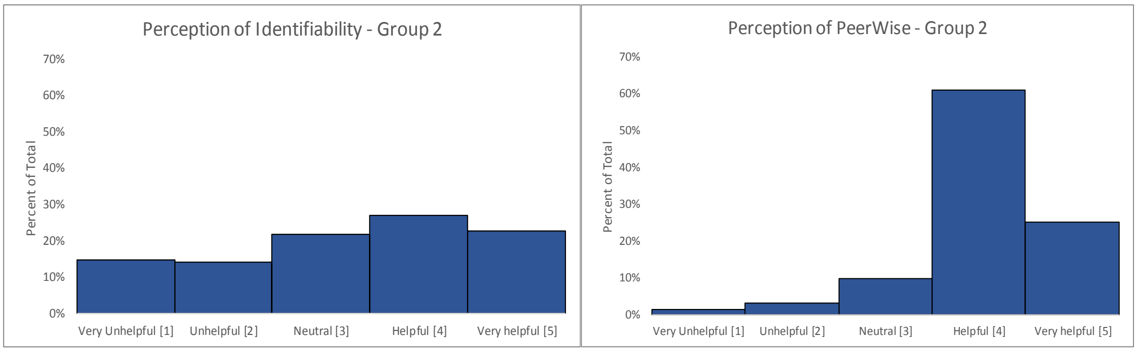


Figure 4.6 - Attitude towards being identifiable and towards the website, Group 2

Attitude towards being identifiable and towards the website: is there a correlation?

Correlation analysis and ordinal regression, were conducted to uncover if the way users felt about being identifiable had any effect on how helpful they considered the website. First, Spearman's rho was obtained, and the correlation coefficient was of 0.128, showing a very weak correlation between the two attitudes. Second, after conducting the ordinal regression, the model was not significant (p=.282), and the Cox and Snell pseudo-R^2 value was just .030, meaning that the attitude towards being identifiable only accounted for 3.0% of the variation in the overall perception of the website. Therefore, it can be concluded that, for students in Group 2, the attitude they had regarding how useful it was to be identifiable had a minimal, essentially uncorrelated, impact on the perception they had towards the website as a whole. Appendix 4.3 presents the full results of these tests.

4.4 SUMMARY OF FINDINGS

As outlined in Chapter 3, the first stage of the quasi-experiment focused on the comparison of a CC and a SNS regarding self-presentation. Namely, the first research question of this thesis investigated the effects of anonymous and identifiable users in online environments. Hence, participants in Group 1 were presented solely with content that could be rated, whereas for users in the second cohort the content was accompanied by the identities of those who authored it.

The first research question comprised three objectives. The first one to be tackled, OB-3, consisted of investigating if – and how – being identifiable might affect the transmission of UGC, in comparison to being anonymous. Results showed that the choices (i.e. ratings) made by anonymised users differed significantly from those made by users who were identifiable. Further, ratings were not only different regarding average values, but also in the way in which the rating scale was used: in the case of anonymous participants, they used the full range of the scale, and ‘fair' (2) was the most used rating. In contrast, identifiable users avoided extreme values, in particular 0, 1 and 5, and rated mostly with ‘good' (3).

From the viewpoint of Tversky & Kahneman’s studies of the psychology of choice (1981), it could be said that changing the website's design had an impact on the judgement of users. That is, people who were presented with anonymised information and those who interacted with their real identities faced a different decision-problem, and therefore adopted a different frame, which produced different choices. Moreover, from social media classification (Kaplan & Haenlein 2010) it can be argued that, even when presented with the same UGC, users from SNSs and CCs would make different choices due to the different levels self-presentation that characterises them. Consequently, online users might rate UGC differently when their names and other personal characteristics are displayed together with their reviews, especially if they know that real-life acquaintances can identify them. Thus, linking this argument with the introductory section of this chapter regarding identity, it could be argued that Goffman’s (1959) theory about how the ‘self’ varies depending on the ‘audience’ observing one’s ‘performance’ still holds for online environments. This point will be further explained in the theoretical discussion of the thesis, presented in Chapter 8. Furthermore, the findings from this chapter also suggest that when people search for advice in online environments, they might be prone to selecting different recommendations depending on whether they browse a CC or a SNS.

The second addressed objective (OB-7) consisted in evaluating which group of biases had a stronger effect on the selection of information. When only considering content and prestige, it was found that prestige-based variables had the strongest effect on ratings for participants in the first cohort. Specifically 'online-gained prestige' (i.e. badges) was the variable with the greatest impact, suggesting that gamification (e.g. Denny 2013; Robson et al. 2016) plays an important role when users are anonymous. Conversely, in Group 2, the variables with the strongest effect on ratings appeared to be content-based, although the comments from students indicate that personal relations might have played an important role. However, personal relationships will only be introduced in Chapter 5.

Moreover, while investigating biases and testing the effect of each independent variable, six emerging findings were encountered, two related to content and four concerning prestige. Firstly, regarding content, the length of the posted content (i.e. a question's number of characters) and the presence of a web link had a positive, significant effect on ratings for both cohorts. This is consistent with research performed in respect of online reviews showing that reviews which are longer and include pictures are perceived as being of better quality (Liu & Park 2015; Cheng & Ho 2015; Park & Nicolau 2015). Secondly, the date when the questions were posted and rated also affected ratings. However, the date when questions were posted was found to have opposite effects when students were anonymous compared with when they were identifiable; the former appeared to elicit a 'competitive' strategy, while the latter seemed to have produced a 'collaborative' orientation.

Thirdly, concerning prestige, it was found that the mechanisms PeerWise has in place to give 'online-gained prestige' to users (e.g. badges) were significant for both groups, but had the greatest effect when students were anonymous. Fourthly, it was found that questions posted by staff (classified as 'real-life prestige') were treated differently in each group. When students were anonymous and only the module staff could recognise them, questions posted by the latter were rated higher than average. Conversely, when ratings and identities were known to the whole group, students evaluated the questions of the module team below the group's average. Thus, this fourth finding contradicts the overall trend of the groups, where anonymity produced 'harsher' ratings than when students were identified. A plausible explanation has to do with identities (e.g. Goffman 1959; Reicher et al. 1995) and the presence of ingroups and outgroups (e.g. Tajfel & Turner 1979), a point elaborated on in the Discussion (Chapter 8).

Fifthly, when comparing the percentage of targeted questions between distinct nationalities, it was found that almost all decreased when students were identifiable. Specifically, it was discovered that students from the UK and other nationalities decreased the number of answers towards questions authored by Chinese students; yet students from China increased the percentage of UK-targeted questions. These results were similar to those obtained by Wu et al. (2011) who studied English-based tweets from many different nations and found, among other things, that Indians would follow and share significantly more content from Americans than *vice versa*. They argued that the perceived power of one nation over the other was reflected in the interactions of online users. Hence, given that Chinese students are enrolled in the UK education system (and not *vice versa*), when identifiable, Chinese might have been more inclined towards accessing the UGC posted by their British peers.

Sixth and last, gender was not a significant variable in explaining ratings when users were anonymous, but became significant when they were identifiable, with ratings for females being lower. This finding is in line with many fields where only by submitting something with a woman's name, instead of a man, the information is perceived as having less quality or truthfulness (e.g. Handelsman & Moss-Racusin 2012). It should be noted that although gender and nationalities raise some interesting points for debate and reflection, they are not the focus of this thesis, and will not be discussed in-depth. Nonetheless, further research should evaluate these issues further, as they might be of relevance, especially for online education.

Finally, the chapter's last objective, OB-4, was to determine whether the attitude of users towards being identifiable affected their overall views of the utility of the website. Regarding this matter, it was found that being identifiable had a very weak correlation with the overall perception about PeerWise. Strikingly, two-thirds of respondents to the questionnaire administered to Group 2 preferred that they be signed-in anonymously, stating this would make ratings and other online interactions fairer and reduce the potential for biased ratings. Conversely, the third that declared a preference towards being identifiable claimed that this enhanced the quality of the posted content and could prevent bullying.

CHAPTER 5: THE EFFECT OF DIFFERENT LEVELS OF SELF-DISCLOSURE ON CHOICES

As explained in Chapter 2, both self-presentation and self-disclosure are part of identity management. Self-disclosure was defined as the revelation of personal information that is aligned with the image online users want to give about themselves (Kaplan & Haenlein 2010). It involves information about the users' emotions, attitudes, feelings, thoughts, likes and dislikes (Moon 2000; Kaplan & Haenlein 2010). For this study, self-disclosure is studied through the likes and dislikes that users express through the available rating scales for different websites.

Ratings reflect the quality of the ideas posted in online environments, and rating scales are used to assess almost every imaginable category, such as videos, movies, consumer electronics, travel services, teachers, coding, and books (Riedl et al. 2010). Hence, ratings serve as online feedback mechanisms that disseminate word-of-mouth information about experiences, products and services among networks (Dellarocas 2003). Unfortunately, rating scales are under-studied, and there are no clear guidelines regarding their design (Riedl et al. 2010; Riedl et al. 2013). On the one hand, scholars tend to suggest that longer scales are needed (e.g. Riedl et al. 2010; Riedl et al. 2013), whereas practitioners appear to prefer using shorter rating systems (e.g. YouTube 2009; Ciancutti 2011).

In the absence of guidelines regarding the design of rating scales, some websites change them without understanding the consequences for the transmission of UGC or the choices of users. To name a few examples, in 2010 YouTube, the video sharing site, changed its five-point rating scale for a dichotomous one. As explained in Chapter 2 (see Figure 2.3), this happened because users would only sometimes rate videos when they did not like them but overwhelmingly rated them when they really enjoyed them, hence ignoring the in-between scores (YouTube 2009). Moreover, in 2016, Facebook – the biggest SNS in the world – introduced a range of 'reactions' in addition to its 'like' button (Krug 2016; Teehan 2016). This change took place after twelve years of only allowing a single-rating button and receiving criticism from their users for not giving more options to respond to the content of users. Other recent examples are Spotify, which in 2012 introduced a thumbs up/down rating, and Netflix which around the same time experimented with half and full five-point likert scales, and finally adopted a dichotomous scale in 2017 (Smith 2017). Therefore, it can be seen that both SNSs and CCs have made changes through trial and error in order to decide which rating scale is better for understanding their customers and for their users to evaluate each other's content.

5.1 SECOND RESEARCH QUESTION AND OBJECTIVES

This chapter therefore addresses the second research question of the thesis: *How are the choices of users affected by different levels of self-disclosure happening due to the use of distinct rating scales?*

To address this research question, the chapter focuses on tackling objectives 5, 6, and 7, outlined below. As with the previous chapter, the objectives of the present one remain the same as in Table 2.4. Figure 4.1 illustrates how these objectives fit with the conceptual model of the thesis.

- **OB-5:** To investigate if – and how – different rating scales affect the choices of users, by comparing two types: likert and dichotomous.
- **OB-6:** To determine if the use of different scales has: 1) any social implications when users are identifiable, and 2) any effect on the attitude towards the website's design.
- **OB-7:** To detect how the three group-biases take place in online environments, focusing on the study of conformity, which has been poorly explored both offline and online.

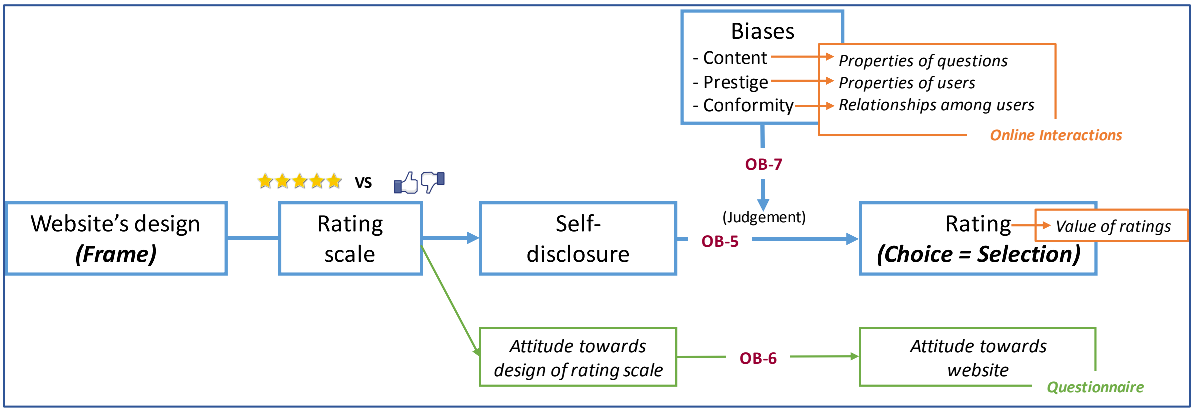


Figure 5.1 - Objectives of second research question and link with conceptual model

5.2 METHOD

5.2.1 Participants

As previously mentioned, all participants were third-year undergraduate students. The total population used in this chapter is of 678 students, 351 females and 327 males. Group 2, where participants used the likert scale, comprised 369 pupils, 330 based in the UK and 39 in China. Further, 309 users; 278 based in the UK and 31 in China, experienced the dichotomous scale (Group 3).

5.2.2 Quasi-experimental set-up and data collection

In Group 2, students could rate the questions posted by their peers with a likert scale, in which values went from zero ('very poor') to five ('excellent'). Conversely, in Group 3 the rating scale was changed to be dichotomous, and students could only rate questions using 'like' and 'dislike'. To facilitate the comparison among years and to avoid affecting other aspects of the webpage[26], 'dislike' was given a numerical value of zero and 'like' a value of five points. Figure 5.2 shows how the two rating scales were designed. Appendix 3.1 contains numerous screenshots showing how PeerWise was set up for these cohorts.

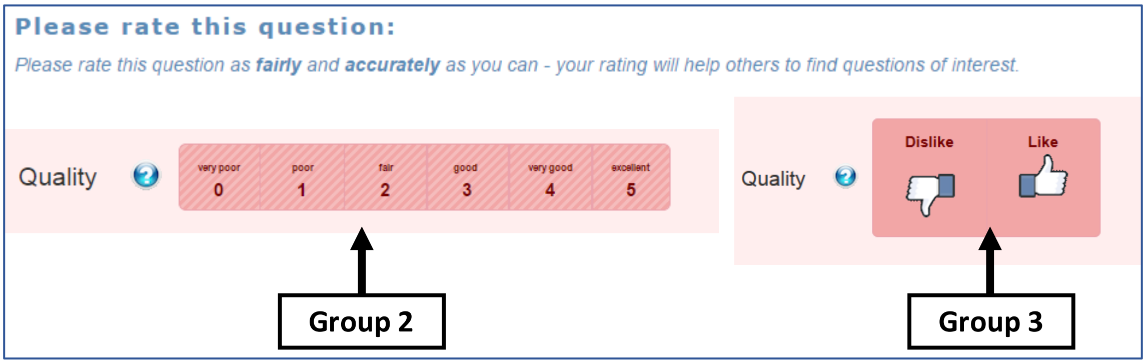


Figure 5.2 - Rating scales used on Groups 2 and 3

Regarding data collection, online interactions were extracted from PeerWise, and later compared using SPSS. In addition, participants in both cohorts took part in a questionnaire (N=392) and six focus groups (N=29), three per group. For this chapter, the quantitative data from the survey were analysed using Excel and SPSS, and the qualitative

[26] As mentioned, PeerWise had many features like badges and reputation scores, which were designed to work with values between zero and five. Further, the average scores per question are designed to display a value between zero and five.

data coming from the open-ended questions of the survey and the focus groups were analysed with Excel and NVivo.

5.2.3 Procedure

The procedure for data analysis is very similar to that of Chapter 4, so the discussion here highlights the points where they differ, with only the new statistical tests and variables described in detail.

A. *Investigating the effects of two rating scales on ratings (Figure 5.1, OB-5)*

First, to have an overview of the primary online interactions of both groups, the average and maximum values were contrasted. Second, a between-group comparison of the ratings was performed. Comparing rating scales with different numbers of values is not straightforward and there is no 'correct' way to do this; in one sense, it is like asking two different questions. Accordingly, three comparisons were performed: 1) Using the 'original' (raw) values. As mentioned, in both groups the minimum rating value was zero, and the maximum was five; hence, all ratings were compiled using the same range. 2) Aggregating at the question-level. Given that both groups had the same range of values [0-5], it was also possible to aggregate ratings at the question-level (i.e. obtain the average rating per question), and then perform an independent sample t-test. 3) Transforming values from the likert scale to dichotomous. Again, there was a number of ways in which this could have been done, but the chosen transformation was the one which, arguably, made more sense mathematically. As the likert scale comprised six values [0-5], those of [0, 1, 2] were considered 'dislike' and the values of [3, 4, 5] were transformed into 'like'. It should be noted that the scale for Group 3 was binary, and therefore it cannot form a normal distribution. For this reason, options 1 and 3 described above, needed tests to be non-parametric and categorical, and hence a *chi-square* test was conducted in these cases (Field 2013).

Third, after conducting the between-group comparison for all ratings, the analysis was repeated for the 20 identical questions posted in both Groups 2 and 3 (see Figure 3.3). Fourth, the findings from the statistical tests were triangulated with the thematic analysis, which focused on investigating why in Group 3 there was a disproportionate amount of 'likes' in comparison to 'dislikes'. As has been explained, the creation of themes followed the six phases of conducting thematic analysis suggested by Braun and Clarke (2008).

B. *Exploring which group of biases had the strongest effect on ratings (Figure 5.1, OB-7)*

This chapter analyses the three group-biases: content, prestige, and conformity. Chapter 4 has already listed and defined all the independent variables considered content-based or prestige-based. Thus, only conformity-based variables are outlined below.

Conformity-based biases

- *Conformity to the outgroup, i.e. absent ties (Categorical-Binary):* This type of conformity was seen as the ratings given to the online users that were not part of one's personal network. That is, neither the 'author' of the question nor the user who performed the rating (i.e. 'rater') declared knowing the other.
- *Conformity to one's personal network*[27]*, i.e. ingroup (Categorical-Binary):* Conformity to one's personal network is when the 'rater' considered the 'author' to be within her personal network. That is, at the very least, the 'author' and the 'rater' were acquaintances.

As mentioned in Chapter 3, personal networks were detected from the questionnaires sent to students in Groups 2 and 3. On these questionnaires, students were asked to name at least three, but preferably more, students that they knew from their class. It should be noted that this chapter only distinguishes between 'non-friends' (i.e. outgroups) and 'friends' (i.e. ingroups or personal networks) because the focus is firstly on rating scales and secondly on biases. However, the following chapter – which is dedicated to the study of personal networks – makes a further differentiation among different friendship levels, or strengths of ties (e.g. Granovetter 1973).

This chapter follows a similar procedure to Chapter 4. First, Groups 2 and 3 are compared using histograms and standardised coefficients. It should be noted that, to conduct a regression, all variables were aggregated at the question level. Therefore, variables that

[27] Personal networks comprised the following strengths of ties: strong, intermediate, weak and N/A. However, these will only be differentiated in Chapter 6.

related to the properties of the question or the author of the question remained unchanged; however, those that regarded the relationship of the 'author' and the 'rater' became a proportion. So, for instance, if a question was rated 20 times, and 10 of these ratings were among the author's personal network, by aggregating at the question level the value of the variable 'personal network' (i.e. the proportion of friends) would equal 0.5. Second, the key similarities and differences among Groups 2 and 3 are highlighted. Third, given that for both cohorts the variable with the strongest impact on ratings was the proportion of friends (i.e. personal networks), it was decided to undertake a further test to better understand how the rating scale interacted with friendships across groups. This was performed by making use of the *three-way chi-square*, also known as *log linear analysis* (Field 2013) or *three-way contingency table* (Lowry 2001).

C. Determining the effects of the attitude towards the rating scale on the attitude towards the website (Figure 5.1, OB-6)

Similar to Chapter 4, to test the effect of the attitude towards being identifiable on that towards the website, two analyses were conducted. First, Spearman's rho was obtained, and then an ordinal regression was performed. These two analyses were performed for each cohort.

5.3 RESULTS

First, both groups are compared regarding all their online interactions, focusing on ratings and their distributions. Second, the three groups of biases are ranked within each group, in order to determine which was the strongest. Third, the attitude towards the rating scales of both cohorts is compared, together with overall perceptions of the website and the effect that the different scales produced on this.

5.3.1 Comparing the impact that different levels of self-disclosure have on choices

Summary of online interactions of Group 2 and Group 3

Table 5.1 presents a summary of the most relevant online interactions between both groups. As can be seen, although Group 3 had 16.3% fewer students than Group 2, students in both cohorts authored the same number of questions, on average. However, the average number of answers, ratings, and comments per student were all substantially lower for the third cohort.

Table 5.1 – Summary and comparison of online interactions between Groups 2 and 3

Online Interactions	Group 2	Group 3
Students (active in PeerWise) **Based in the UK** **Based in China**	Total: 369 UK: 330 CH: 39	Total: 309 UK: 278 CH: 31
Total no. of questions (non-deleted)	2,525	2,125
Max. no. of questions per student	19	28
Ave. no. of questions per student	6.8	6.8
Total no. of answers	34,379	21,486
Max. no. of answers per student	644	596
Ave. no. of answers per student	96.0	72.6
Total no. of ratings	25,976	14,997
Ave. no. of ratings per student	70.4	48.5
Mean rating of all questions	2.94	4.69
Comments (including replies)	13,805	3,314
Max. no. of comments per student	485	81
Ave. no. of comments per student	38.4	11.4
Answered but not rated	8,403 = 24.4%	6,489 = 30.2%
Answered but not commented	20,574 = 59.8%	18,172 = 84.6%
Rated but not commented	12,171 = 46.9%	11,683 = 77.9%
Max. no. of total badges per student	363	162
Ave. no. of total badges per student	48.8	33.9
Max. no. of distinct badges per student	25	25
Ave. no. of distinct badges per student	16.8	15.0

Regarding the primary focus of this thesis – ratings – there was a 5.8% increase in the number of questions that were answered but not rated from Group 2 to 3. Hence, this may suggest that there was some level of discomfort by using the rating scale, given that: ratings were optional for both cohorts, both sets of participants were identifiable, and everything in PeerWise and the module remained the same; except for the rating scales. Moreover, as can be seen even from the mean of the ratings, in Group 3 questions had an average rating of 4.69, 1.75 above the mean of Group 2. Therefore, this shows that students not only rated less often, but also rated using disproportionately more 'likes' (5) than ‘dislikes’ (0). The avoidance of giving negative ratings seen with the change from anonymous to identifiable noted in the preceding chapter, seems to have been accentuated with the use of a dichotomous scale.

<u>Between-group comparison of all ratings</u>

As mentioned in the methods section of this chapter, there are three ways in which ratings can be compared: 1) with the original rating values, 2) by converting half of the likert scale to ‘like’ and the other half to ‘dislike’, and 3) by aggregating ratings to the question level.

1) *Original rating values*

Figure 5.3, below, shows the comparison of the ‘raw’ values from both cohorts. As noted above, the ratings in Group 3 were predominantly five’s (i.e. ‘likes’).

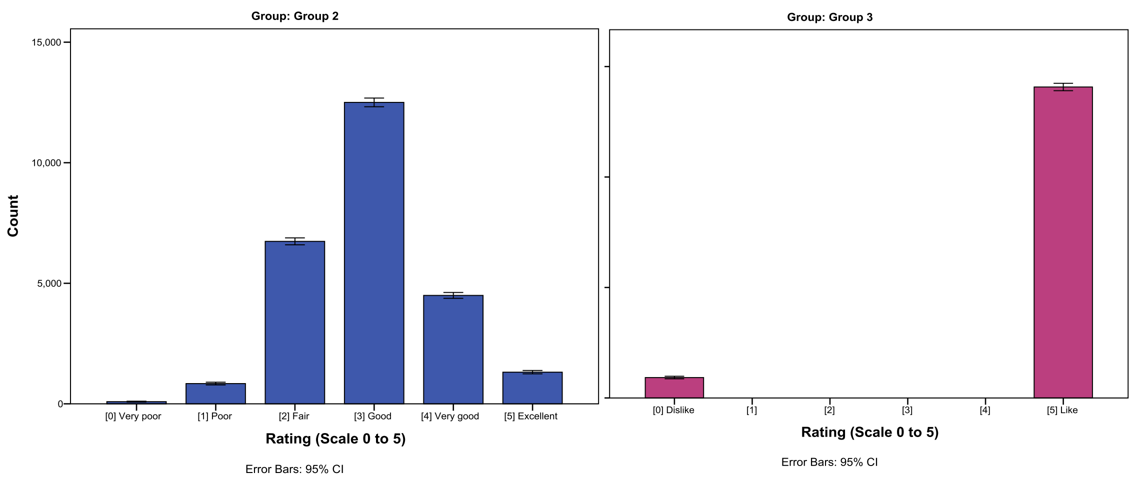


Figure 5.3 - Comparison of all ratings with original rating scales, Groups 2 and 3

Group 2 had an average rating of M=2.94 (SE=.006), while Group 3 had a mean of M=4.69 (SE=.010). However, as the values in the third group were binary, ratings need to be compared by frequency. Thus, Table 5.2 shows the frequency and percentages for each of the values within the scales.

Table 5.2 - Frequency of all ratings, original comparison, Groups 2 and 3

	Value	Frequency	Percent	Cumulative percent
Group 2	[0] Very poor	89	0.3%	0.3%
	[1] Poor	839	3.2%	3.6%
	[2] Fair	6,738	25.9%	29.5%
	[3] Good	12,502	48.1%	77.6%
	[4] Very good	4,496	17.3%	94.9%
	[5] Excellent	1,312	5.1%	100.0%
	Total	***25,976***	***100.0%***	
Group 3	[0] Dislike	923	6.2%	6.2%
	[5] Like	14,074	93.8%	100.0%
	Total	***14,997***	***100.0%***	

As can be noticed in Table 5.2, the total number of ratings with the value 3 (Good) in Group 2 are comparable to the total number of ratings of 'Like' (5) in Group 3; although this statement does not hold true for the percentage values. Moreover, as expected from the average value of the ratings from Group 3, the number of likes are disproportionally large in comparison to the number of dislikes: a ratio of 15.24 to 1, to be precise. Further, given that it is not possible to perform a t-test with this dataset, an independent-samples Mann-Whitney non-parametric test was performed. The test indicates that ratings from both cohorts were significantly different, $U = 356{,}394{,}754$, $z = 146.67$, $p < .001$.

2) *Converting half of the likert scale to like and the other half to dislike*

As explained, the most obvious way in which the scales could be compared, at least from a mathematical viewpoint (given that one scale had six values and the other one two), was to convert half of the values of the likert scale to 'like' and the other half to 'dislike'. Therefore, the values of [0, 1, 2] became [dislike = 0 points], and the values of [3, 4, 5] were transformed to [like = 5 points]. The comparison can be observed in Figure 5.4.

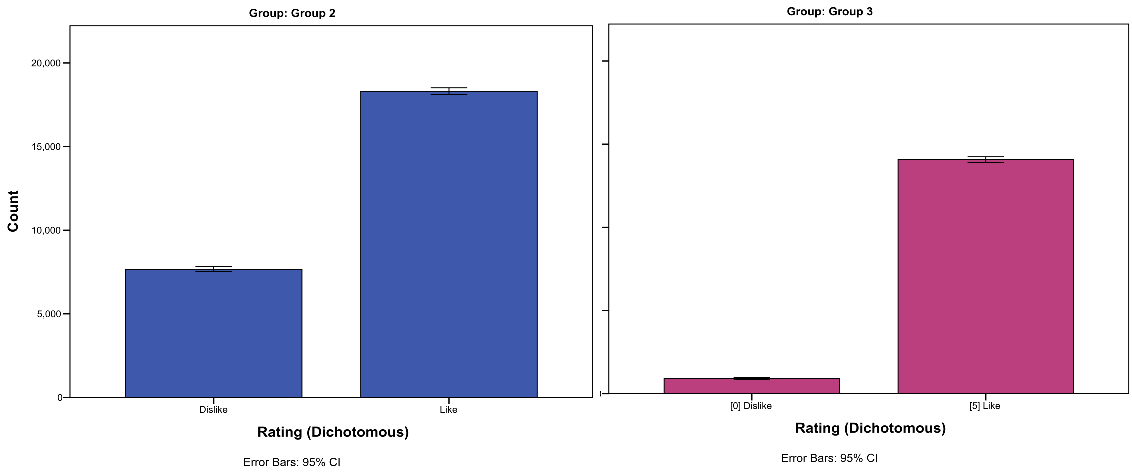


Figure 5.4 - Comparison of all ratings converted rating scales (dichotomous), Groups 2 and 3

With this comparison, 'dislikes' accounted for 29.5% in Group 2 and 6.2% in Group 3, while 'likes' represented 70.5% of all ratings in Group 2 and 93.8% in Group 3. Given that both scales are now binary, a chi-square test was performed, which turned out to be significant at $X^2(1, N=40{,}973) = 3{,}130.74$, $p < .001$.

3) Aggregating ratings to the question-level

Finally, given that both scales ranged from 0 to 5, another way of comparing them was by aggregating ratings at the question-level. This comparison is presented in Figure 5.5. As can be seen, the ratings of Group 3 were not normally distributed, due to the excessive number of questions that were rated purely with 'like' (5). However, it is worth visualising the comparison, as the ranking of variables was performed by aggregating at the question level.

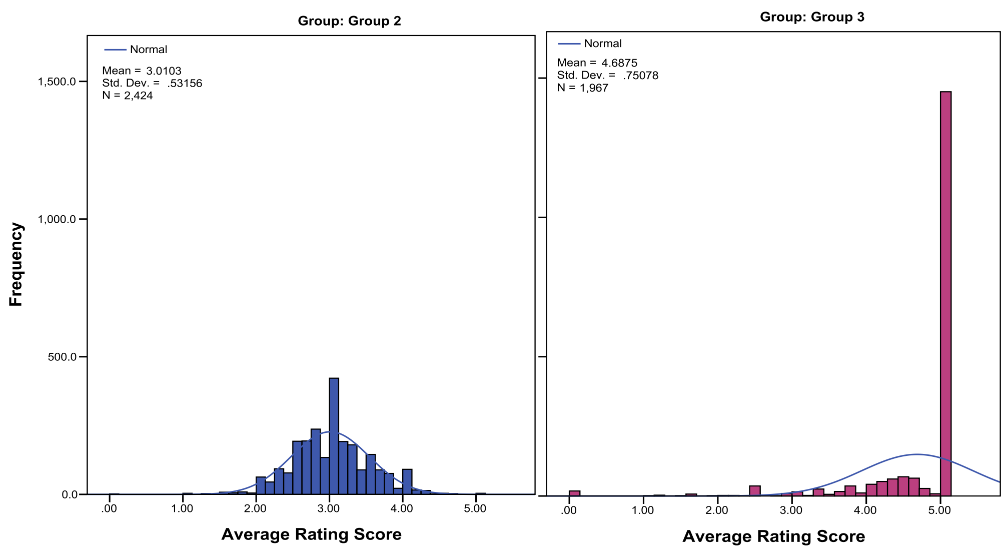


Figure 5.5 - Comparison of all ratings aggregated at the question level, Groups 2 and 3

Between-group comparison of 20 identical questions

The 20 identical questions that appeared in both groups were compared 1) with the original ratings and 2) converting the ratings made with the likert scale into dichotomous.

1) *Original rating values*

When compared with the likert scale, the 20 repeated questions in Group 2 had a mean of 3.08 (*SE*=.04, *N*=546) while those from Group 3 had an average of 4.69 (*SE*=.07, *N*=311). Once again, the average difference was tested using a Mann-Whitney test and was significant at U = 153,620, z = 20.65, p < .001. Figure 5.6 shows how the ratings differed.

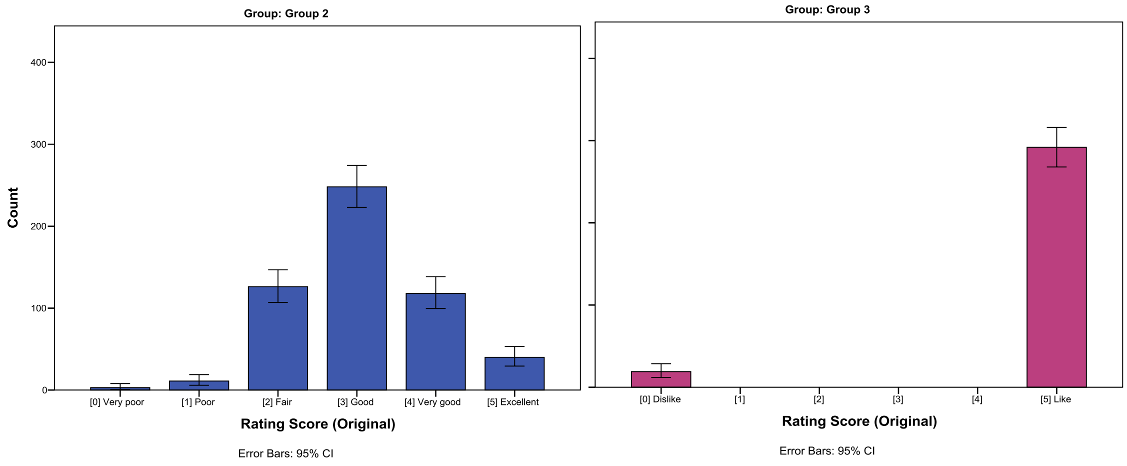


Figure 5.6 - Comparison of 20 identical ratings with original rating scales, Groups 2 and 3

2) *Converting half of the likert scale to like and the other half to dislike*

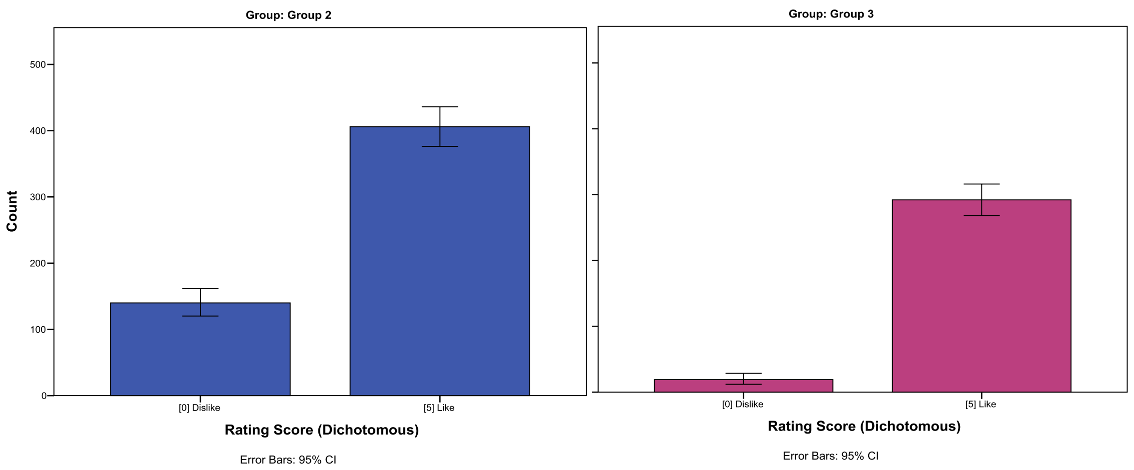


Figure 5.7 - Comparison of 20 identical questions, converting rating scales (dichotomous), Groups 2 and 3

Three differences should be highlighted from the comparison of the 20 identical questions presented in Figure 5.6 and Figure 5.7. First, although the number of questions was the same (20), Group 3 provided 235 (43.0%) fewer ratings than Group 2. This suggests that, for some reason, students preferred not to rate rather than having to choose between giving a 'like' or 'dislike' rating. Second, from all the 20 identical questions, only one had a lower rating average in Group 3; the other 19 questions got higher rating scores with the dichotomous scale. The disproportionate use of 'likes' over 'dislikes' in Group 3 can be appreciated in both the figures shown above.

Third, in addition to the previous point, 9 of the 19 questions that received higher rating averages had an average of 5, meaning that they received *only* 'likes'. It could be argued that these questions are simply of exceptional quality. However, students from the previous two cohorts had also rated them, and the respective averages were notably lower. For example, Question 1 (authored by the module leader) was rated in Group 1 by 35 students and its average rating was of 2.74; in Group 2, 51 students rated the same question with an average of 2.80; and in Group 3 the same question was rated by 35 students receiving an average of 5.00. This finding was surprising, suggesting students in Group 3 were not very critical when assessing information, even though they were told that the ratings within PeerWise did not count towards their grades. Hence, this suggests that the adopted scales affected the transmission of information, as the ratings of participants varied significantly.

Triangulating findings regarding interacting online with the use of the dichotomous scale

The thematic analysis focused on trying to reveal why such a dramatic change occurred in the ratings of users of the third cohort. Specifically, as can be seen in Table 5.3, it seemed as if the dichotomous scale was in part responsible for the lack of ratings from participants and for them avoiding the 'dislike':

Table 5.3 - Thematic analysis regarding the avoidance of 'dislikes' when using the dichotomous rating scale

Theme	Original quotes from participants *[sic.]*
Dichotomous scale perceived as emotional	*"I think having a rating scale from 1 to 5 will be more effective and rational to evaluate one's question"* [sic.] (Quest-G3, inji682).
Difficult choice	*"Sometimes, it is hard to simply decide like or dislike. And a rating scale can help make a ranking list"* [sic.] (Quest-G3, yisu483).
Felt obligated to like	*"I would change it to to a numbering system that than like or dislike just to make them more accurate. Users will feel less obligated to give a like and their ratings will be truer to their feelings when answering questions"* [sic.] (Quest-G3, moma912).
Dislike seemed unkind	*"(I would vote for changing the rating scale to a broader one) Because it's such a small scale and dislike seems a bit too harsh"* [sic.] (Quest-G3, elya294).
Identity-related[28]	*"Having just two options may sway an individual to select 'like' due to the name appearing on the dislike column. Either making the rating anonymous or including a 1-5 may make it more accurate"* [sic.] (Quest-G3, ukki261).
Personal networks	*"Even if it wasn't a particularly good question, I'd still like it cause they were friends [laughs]. Again, that's why I think a helpful scale would probably be more useful"* [sic.] (FG2-G3, raed974).
Avoidance of rating	*"The current system is too black and white and can discourage people from rating average questions that are neither exceptional nor poor"* [sic.] (Quest-G3, myve209).

It should also be highlighted that both in the questionnaire and focus groups of Group 3 students said they would rather not even answer a question if they did not like it: *"Usually every question I would like as I wouldn't answer the question if I disliked it"* (Quest-G3, ahmc850). Moreover, although answers are not the main focus of study, it is important to mention them, as they were a previous step to rating and an important part of the selection

[28] Combined effect of identifiable user profile and dichotomous rating scale.

process, as argued in Chapter 4. The following conversation between two participants that attended one of the conducted focus groups further explains this point (*FG2-G3):*

- Researcher: *“So, even if you came across a question you thought it was [bad quality] you'd rather not rate it?”*
- raed974: *"Yes. Even if I didn't like a question, I wouldn't have clicked dislike, I would have just not clicked like"*
- amca620: *"Unless it was obviously wrong and they'd made a mistake, which would draw their attention to it.*
- raed974: *"But then even, I wouldn't have put a dislike, I would just leave a comment 'I think this is wrong'".*
- amca620: *"Yeah, right".*

In addition, it came as a surprise that a number of students suggested that those who disliked questions should be forced to give a comment explaining why they were doing this. However, there is an interesting asymmetry here in that no one said this about liking, suggesting that those who got a dislike felt wrongfully evaluated. For instance, “*2 options aren't enough, and if you dislike you should be required to give a reason*” (Quest-G3, ysth385).

To sum up, the tendency observed in the previous chapter where, by being identifiable, students avoided extreme values was exacerbated by the use of the dichotomous scale. Apart from the observations that show a disproportionate use of ‘like’, in the questionnaires and focus groups, students declared not having used the ‘dislike’ rating because it was ‘harsh’ and also because it had real-life implications, an issue that is addressed in the following section.

5.3.2 Impact of content, prestige, and conformity based biases on choices

Using standardised coefficients to rank the effect of the variables on ratings

As in the previous chapter, this one also makes use of linear regression and standardised coefficients to rank the effect of variables within each cohort. However, as the rating scales are different, in order to make a ‘fair’ comparison the regression was performed by aggregating ratings at the question level (see Figure 5.5). Therefore, the variable 'personal networks' appears as a proportion (i.e. proportion of 'friends' who rated the question).

Chapter 6 studies personal networks in greater depth but, to give some background on the 'friendships' variable, these are briefly described here. From the 25,491 ratings among students, 8.03% were among 'friends' (i.e. personal network), and 91.97% were amongst 'non-friends' (i.e. absent ties). Moreover, regarding Group 3, 13.37% of 14,686 total ratings were between 'friends' and 86.63% among 'non-friends'.

Regarding the regressions, in Group 2, 20.0% of the variation was explained by the linear relationship between ratings and the predicting variables, and this was highly significant ($F(10, 2398) = 59.95$, $p < .001$). In contrast, only 2.6% of the variation in Group 3 was explained by the linear relationship between ratings and the predicting variables, although this was again strongly significant at $F(6, 1960) = 8.63$, $p < .001$. Table 5.4 above presents the ranking of variables, using the absolute values of the standardised coefficients. Further, Appendix 5.1 presents the full set of results for the regressions of Groups 2 and 3. It should again be noted that regression models were only used to determine the standardised coefficients rather than to obtain a model that predicts ratings, or to determine if the 'ideal' model should be linear or logarithmic.

Table 5.4 - Ranking of variables using standardised coefficients (Groups 2 and 3)

	Group 2		**Group 3**	
Rank	**Independent variable**	**Std.Coeff. (|Abs|)**	**Independent variable**	**Std.Coeff. (|Abs|)**
1	Prop. of friends (Prs. Ntwk)	.219	Prop. of friends (Prs. Ntwk)	.073
2	Q's number of characters	.206	Question's link	.066
3	Question's link	.130	Question's explanation	.057
4	Author's gender	.116	Q's number of characters	.047
5	Author's distinct badges	.115	Author's distinct badges	.044
6	Total number of ratings	.106	Author's nationality	.039
7	Question's Reference	.078		
8	Q's number of alternatives	.067		
9	Question's explanation	.051		
10	Author's nationality	.049		

As can be observed in the table above, in Group 2 the variable that had the most significant impact was the proportion of 'friends' who had rated the question (i.e. personal networks), which is a conformity-based bias. This was followed by two content-related variables: a question's number of characters and the presence of a link. Thereafter, two prestige-based variables followed: author's gender (real-life prestige) and author's distinct badges (online-gained prestige). It should be noted that in the previous chapter, content-based biases appeared to have the strongest impact on Group 2, given that the impact of friendships had not been introduced. However, the above-shown ranking confirms the findings from the qualitative input of participants: personal biases had the strongest impact. Thus, it can be established that the order of biases for Group 2 was: conformity, followed by content, and lastly prestige.

Moreover, in Group 3 the variable with the most impact was also the proportion of friends (i.e. personal networks), followed by three content-biases, and lastly two prestige-based biases. It should be said that the ranking in Group 3 is shorter than that of Group 2 because many variables were not significant, despite the threshold of this group being of $p<.10$. The reason for this is that, as seen in the histogram of ratings aggregated at the question level (see Figure 5.5), the disproportionate use of the 'like' button resulted in many questions having an average rating of 5, therefore skewing the histogram to the right and affecting the normality of the ratings.

To sum up, the effects of biases in Groups 2 and 3 were very similar. There are two important points that should be highlighted. First, the proportion of friends that rated a question is the variable with the biggest impact on ratings in both groups. This was also reflected in the comments made by participants, for instance: *"I feel that the ratings are more of a reflection of the number of friends you have on the course"* (Quest-G2, camu282). For this reason, the following section focuses on getting a deeper understanding of the effect of rating scales on personal networks. Second, regarding the ranking of group-biases, both cohorts followed the same order: conformity followed by content, and then prestige. Regarding nationality (considered real-life prestige), it is worth mentioning that, in both cohorts, students from the UK received the highest average ratings.

Exploring the effect of the rating scales on personal networks

The previous sections showed that in Groups 2 and 3 the variable that had the strongest effect on ratings was the proportion of friends who rated the question; that is, the author's personal network. However, it is unclear the effect that the rating scale had on ratings performed between ‘friends’. It should be noted that the focus of this chapter is not on friendships *per se*, but to investigate the effect that rating scales had on this variable.

In order to make a comparison, both cohorts were contrasted using a three-way chi-square test. To do this, both sets of ratings were compared in binary form, i.e. the Group 2 rankings were converted into a dichotomous representation, as shown in Figure 5.4; namely [0, 1, 2] = dislike (0), and [3, 4, 5] = like (5). To give an overview of the contrasting ratings given to ‘friends’ versus ‘non-friends’, the mean rating of Group 2 was of M=2.94 (SE=.006); the average rating among friends was of 3.61, while among non-friends was of 2.88. Likewise, Group 3 had a mean of M=4.69 (SE=.010); the average rating among friends was of 4.89, and among non-friends was of 4.66.

Moreover, as can be observed in Figure 5.8, in neither cohort did it matter much if the ‘author’ and ‘rater’ were friends or not, in order to get a ‘like’. However, being ‘friends’ did matter for ‘dislikes’, as only non-friends seemed to get them. This trend was more evident for students using the dichotomous scale in Group 3. The effect of ‘dislike’, per group, on the ratings of friends and non-friends was significant at $X^2(1, N=8,430) = 20.48$, $p < .001$. Likewise, the effect of ‘like’, per group, on the ratings of friends and non-friends was significant at $X^2(1, N=31,747) = 200.68$, $p < .001$. The complete set of chi-square tests can be found in Appendix 5.2.

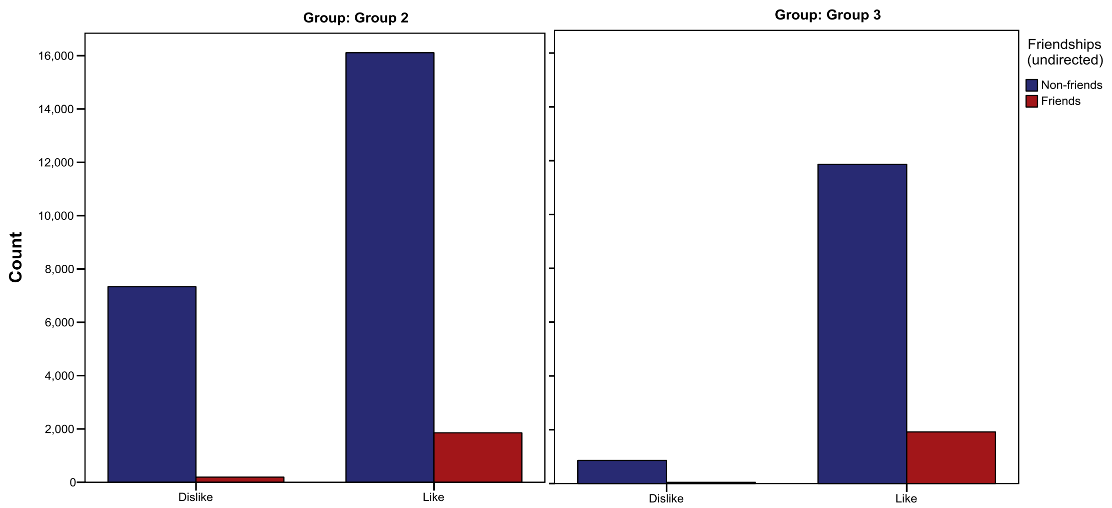


Figure 5.8 - Three-way chi-square of the effect of rating scale on ratings among ‘friends’ and ‘non-friends’

The tendency for participants using the dichotomous scale to avoid disliking questions from their real-life personal networks was further confirmed by the qualitative data from questionnaires and focus groups. Specifically, during one of the focus groups, students gave a very detailed description of how the rating scale had affected their ratings – in particular those regarding their 'friends' *(FG3-G3)*:

- Researcher: But why was it important to rate your friends good if it didn't affected their grade?
- myve209: "I think that criticizing your friends is a lot more difficult than criticizing strangers"
- anse680: "Yeah, it's like a public space, as well. So you can see who's written what comments about who, and that kind of thing. Like, it does have a weird social impact [laughs].
- nehe986: "Maybe the like/dislike was quite strong as well. It would have been better if there was like a number system or something, 'cause just it came across as a bit strong"
- Researcher (follow-up): So what would you suggest?
- Discussion (all): "a number scale, like 1 to 10", "yeah 1 to 10", "or even 1 to 5", "or smiley faces or something" [laughs]
- Researcher (follow-up): So, if it had been a 1 to 10 scale, instead of a like/dislike, would you still have rated your friends' questions higher or just like the rest?
- nehe986: "If it wasn't that good I'd probably not given them a 10, perhaps an 8 or something"
- anse680: "Yeah, even giving them just an 8 instead of a 10 it already says something about the question"

Therefore, both the quantitative and qualitative elements of the research strongly suggest that PeerWise users from Groups 2 and 3 felt pressured to rate their friends higher, but this was heavily accentuated by the use of a dichotomous scale.

5.3.3 Attitude to the two rating scales and their impact on the website's perception

This section discusses users' attitudes towards the rating scale, both in Group 2 and Group 3, and then evaluates the impact this might have on overall attitudes towards the website.

Attitude towards each rating scale: likert versus dichotomous

In the YouTube case presented at the start of the chapter, it was described how the company decided to change their 1-to-5 likert scale to a dichotomous one because users would mostly rate using extreme values (i.e. 1 or 5). Therefore, it was expected that, if a similar situation happened in Groups 1 and 2, participants in Group 3 might find the dichotomous scale more enjoyable. However, in the first two cohorts, students tended to rate mostly with the central values. What is more, in Group 2 – which was set up identically to Group 3 except for the rating scale – the most commonly given rating value was 3 (Good). Hence, it was somehow expected that users from the third group might not enjoy having only two options to rate their peers. Figure 5.9 shows how students from Groups 2 and 3 (*N*=186 and *N*=206, respectively) perceived the helpfulness of each of the rating scales. As can be seen, the majority of users from Group 2 considered the likert scale to be either helpful or very helpful (75.3%). In contrast, this percentage dropped to 55.8% for those users who had to rate using the dichotomous scale.

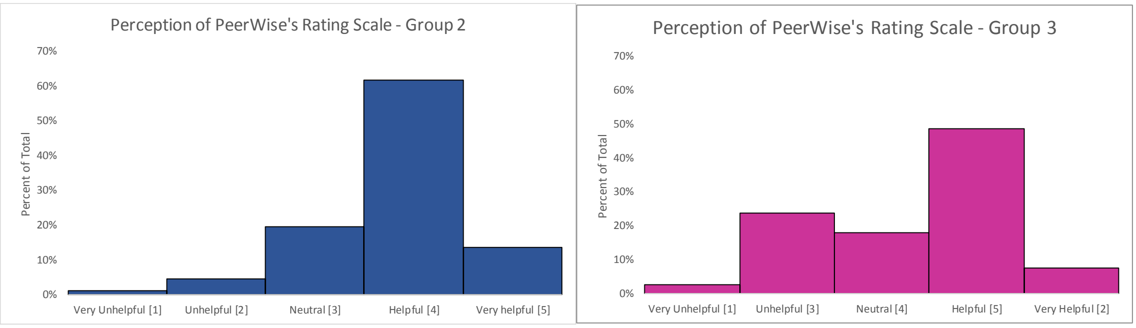


Figure 5.9 - Attitude towards the rating scale, Groups 2 and 3

Furthermore, on the questionnaires of both groups, students were asked if they would: 1) keep using the same rating scale, 2) give more options, or 3) give fewer options. Responses are shown in Table 5.5, together with one quote corresponding to each of the previous scenarios:

Table 5.5 - Question regarding keeping or changing the rating scale, Groups 2 and 3

Group 2: Likert scale *(N=186)*		Group 3: Dichotomous scale *(N=206)*	
Percentage	**Quotes from users [sic.]**	**Percentage**	**Quotes from users [sic.]**
Keep it as it is **55.9%**	"*Feedback from excellent to very poor gives a good indication to the author of how much to improve in future*" [sic.] (Quest-G2, onta342).	**Keep it as it is** **24.8%**	"*ranking in 2 parts it's better so people will decide between helpful or not, rather than go neutral*" [sic.] (Quest-G3, iaba659).
Give more options (i.e. 1-10) **32.8%**	"*Easier to decide what a question is worth in a scale of 1-10*" [sic.] (Quest-G2, iapo175).	**Give more options (i.e. 1-5)** **67.5%**	"*The vast differences in quality of questions perhaps merits more than a binary choice. I would be in favour of the 1 to 5 scale*" [sic.] (Quest-G3, eyfi118).
Give less options (i.e. L/D) **11.3%**	"*There are too many options now. Sometimes others rate my questions as good but I don't know they like my question or not*" [sic.] (Quest-G2, enli050)	**Give less options (i.e. Like)** **7.8%**	"*Keeping it simple and prevent students hit the 'dislike' just because they got the wrong answer*" [sic.] (Quest-G3, king186).

The amount of students that would have kept the rating scale unchanged dropped by 31.1% from the second to the third cohort, while there was a proportional increase (34.7%) in the number of pupils that would have given more options for rating. Hence, this suggests that the majority of users in Group 3 found that the dichotomous scale was somehow restrictive. However, it is interesting to note that in both cohorts about 10% of participants voted for having shorter scales. In the case of Group 3, this would have meant only using a 'like' button, as most SNSs do.

<u>Perceived 'fairness' of ratings with each scale</u>

Participants in both cohorts were also asked to evaluate the fairness of the ratings they received, from mostly unfairly to mostly fairly. Ironically, although the majority of students from Group 3 felt that the dichotomous scale was not helpful, they also perceived that their questions had received fairer ratings than students from Group 2. As can be seen in Figure 5.10, 57.5% of students from Group 2 considered that their questions had been evaluated sometimes or mostly fairly, in contrast with 78.6% of students in Group 3.

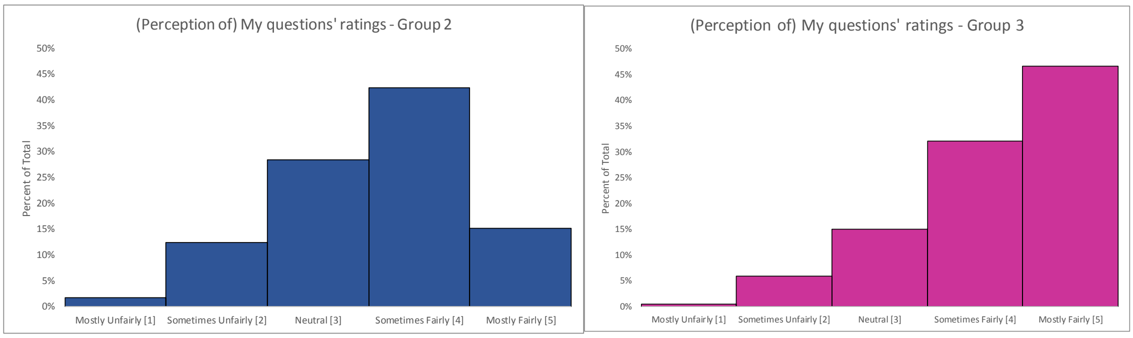


Figure 5.10 – Perception of fairness of given ratings, Groups 2 and 3

Overall attitude towards the website

Moreover, students in both cohorts were asked about their overall perception about PeerWise, and the results were almost identical, with Group 2 having a mean of M=4.06 and Group 3 of M=4.02, in addition to very similar distributions (see Figure 5.11):

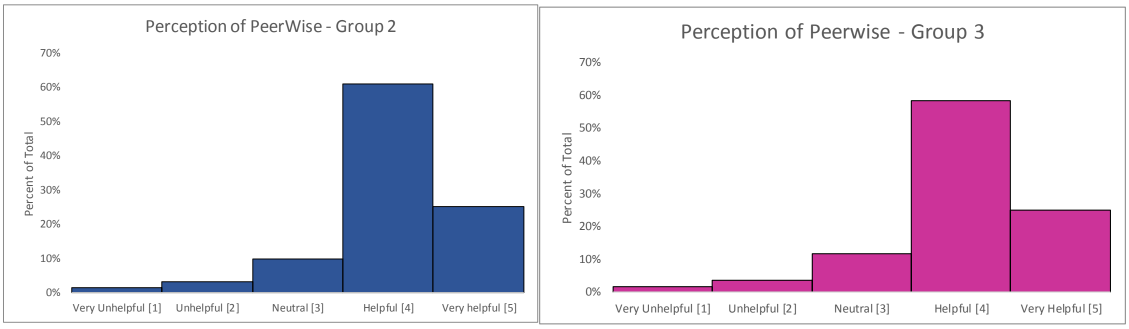


Figure 5.11 – Overall attitude towards PeerWise, Groups 2 and 3

Attitude towards the rating scales and the website: is there a correlation?

Finally, Spearman's rho correlation was calculated together with an ordinal regression, in order to test the relationship between the attitude towards the rating scale and its impact on that of the website.

Regarding Group 2, the correlation coefficient was of .374, showing a mild correlation between the attitude towards the likert scale and that of the website. As evidenced by the Cox and Snell pseudo-R^2 value of .215, p<.001, just over a fifth of the variation in attitudes towards PeerWise were accounted for by attitudes towards the rating scale. Similarly, in Group 3 the correlation coefficient was of .329, and Cox and Snell pseudo-R^2 value was of .214, p<.001, explaining almost the same percentage of the variation in attitudes towards the website. A full set of tests can be found in Appendix 5.3. Therefore, the statistical tests of both cohorts show that there was a moderate correlation between the attitudes towards the rating scale and the website, although this was slightly stronger for Group 3.

5.4 SUMMARY OF FINDINGS

This chapter has aimed to compare social media sites with different degrees of self-disclosure. In order to achieve this, two of the most characteristic rating scales used in CCs and SNSs were compared: likert and dichotomous. Students from Group 2 were allowed to rate multiple choice questions with a 0-to-5 likert scale, whereas those in Group 3 could rate using only ‘like’ and ‘dislike’.

The second research question encompassed three objectives. The first one to be addressed (OB-5) involved the comparison of two rating scales, likert and dichotomous, to investigate if – and how – different scales would affect the transmission of UGC. Three comparisons between these two scales were performed: with their original values, aggregating ratings to the question level, and converting the likert scale into the dichotomous. All of the comparisons showed that there was a significant difference between the two sets of ratings. Specifically, when comparing the 20 identical questions between groups, it was found that the ratings given were significantly different. What is more, about half of these questions obtained an average rating of five (i.e. pure ‘likes’), which was considered unusual given that students were asked to be critical, and these questions had been posted by the ‘Admin’ which had been evaluated below 3 (Good) on average in the previous two cohorts. Hence, it appeared as if students had been presented with completely different content. Therefore, using concepts from Prospect Theory (Tversky & Kahneman 1981), the users from the two quasi-experimental conditions seemed to have experienced two different decision-problems and therefore adopted distinct frames which lead to different choices.

Moreover, a second finding was that people who used the dichotomous scale disproportionately gave ‘likes’ while avoiding the ‘dislike’ button, and this was aggravated when ratings were among ‘friends’ (i.e. personal networks). A plausible explanation is that individuals tend to link their self-esteem to those from their ingroup, therefore benefitting them in a number of ways (Tajfel & Turner 1979; Turner & Reynolds 2012). Further, a third discovery was that, despite having authored the same number of questions per student, those who were required to use the dichotomous scale answered fewer questions, and provided fewer ratings and comments per student. Hence, in terms of the VSR mechanisms (e.g. Campbell 1960; Weick 1969; Aldrich & Ruef 2006; Mesoudi 2011) it can be argued that it was not only selection which was affected, but also variation. All of these findings will be discussed further in Chapter 8.

The second objective to be tackled, OB-7, was to evaluate the three groups of biases and to determine which had the greatest effect on ratings, with a particular focus on conformity. Both cohorts presented a very similar ranking of variables with the most influential group being conformity, followed by content and then prestige. Further, the effect of rating scales was tested on personal networks, and both qualitative and quantitative evidence strongly suggest that participants from both cohorts felt pressured to rate their friends higher, but this was heavily accentuated by the use of a dichotomous scale.

The final objective (OB-6) was to test if users' attitudes towards the rating scale had an impact on their attitude towards the overall website. Results showed that there was a positive and significant correlation between attitudes towards both rating scales and PeerWise. Moreover, there were two relevant emerging findings. First, it was discovered that online users who used the likert scale found it substantially more useful than those who used the dichotomous one. In addition, almost 70% of the students who used the latter and responded to the questionnaire declared that they would have liked to be given more granular options to rate their peers' questions. Second, and in contradiction with the first emerging finding, although participants who used the dichotomous rating schema found it less useful, they also perceived that they had been evaluated in a fairer way than students assigned to the likert scale. Hence, paradoxically, PeerWise users from the third cohort felt discomfited by only having two options to rate questions, yet because this generated 'likes', in turn, students felt a sense of fairness towards evaluations of their questions.

CHAPTER 6: THE EFFECT OF PERSONAL NETWORKS ON CHOICES

In recent years, there has been an increase in social media sites highlighting to their users what others in the network are doing, wearing, buying, and so forth. Further, with the creation of ‘smart lists’ and ‘circles’ from the SNSs of Facebook and Google+ respectively, it has been possible to make users more aware of what their ‘friends’ are reading, buying, liking, and watching, among others. This raises the question: to what extent are users being affected by their friends and other users? Probably this is a difficult question even for the top SNSs because, although they can analyse millions of online interactions, it is tough to determine why a user has liked a specific UGC: because they actually liked it or because they became aware that most of their friends did? In addition, the extent to which users are aware of the influence of their personal network on their choices is open to question.

As explained further in Chapter 2, there is a need for more studies to investigate conformity (Richerson & Boyd 2005), especially in online environments, as most studies that deal with social media and ratings do not take into account the effect that other users have on the choices of individuals (e.g. Riedl et al. 2010; Wu et al. 2011; Wilkinson & Thelwall 2012; Riedl et al. 2013; Liu & Park 2015; Park & Nicolau 2015). Moreover, the way in which conformity has been studied to date has mostly been based on the assumption that all members of a group know each other (e.g. Henrich & Boyd 1998; Henrich 2004; Richerson & Boyd 2005; Perreault et al. 2012). However, although this may be true for some communities, it is far from the reality of online environments. With almost half of the world’s population online, it is not possible that all users know one another. This thesis emphasises the need to distinguish between the whole and personal networks, based on the concept of outgroup and ingroup (Tajfel & Turner 1979). In addition, it proposes a further differentiation within personal networks based on the strength of ties (e.g. Granovetter 1973; 1983).

The fact that online users can differentiate among those that are known to them in real-life might be beneficial given that people evolved by foregoing the costs of individual learning and by relying on others through teaching and imitation (Mesoudi 2011). Hence, by requiring users to sign-in with their real identities, websites allow them to quickly detect those whom they already trust. Conversely, this also allows websites and companies to exploit peoples' personal networks. As mentioned in Chapter 2, some examples of sites that have integrated with SNSs are Quora, TripAdvisor, Netflix, Spotify, Instagram, and Amazon, among others. Thus, CCs that allowed anonymity are now strongly encouraging users to disclose to their networks what they are watching, listening to, and buying. Arguably, the constant presence of others might have an impact on people's identity and, consequently, on the way they access information and make choices.

This chapter aims to determine if the presence of users' real-life personal networks (i.e. ingroups) has an effect in the way users perceive (i.e. access and rate) UGC. Further, this chapter investigates whether there is a difference in how people conform, depending on the strength of their relationships with other users. To achieve this, the ratings given between different levels of relationships were compared. It should be highlighted that, unlike the two previous chapters, some of the methods used in this one are rather exploratory. Therefore, in some cases, the researcher explores different ways to arrive at a solution and then presents the one considered to be most adequate, discussing the advantages, disadvantages, and implications of these methods.

6.1 THIRD RESEARCH QUESTION AND OBJECTIVES

This chapter addresses the third research question of the thesis: *How are the choices of users affected by the online presence of their personal networks?*

To address this research question, the chapter has focused on tackling objectives 8, 9, and 10, outlined below. Further, Figure 6.1 shows how these objectives fit with the conceptual model of the thesis.

- **OB-8:** To determine if – and how – users' real-life ingroups and outgroups affect their choices, and thus the transmission of UGC.

- **OB-9:** To make a further differentiation of ingroups using the strength of ties, and evaluate if these have any effect on choices. Further, to assess if online interactions can predict real-life relationships.
- **OB-10:** To understand: 1) the effect that online interactions might have on relationships and vice-versa, and 2) the extent to which online users are aware of the impact that others have on their choices.

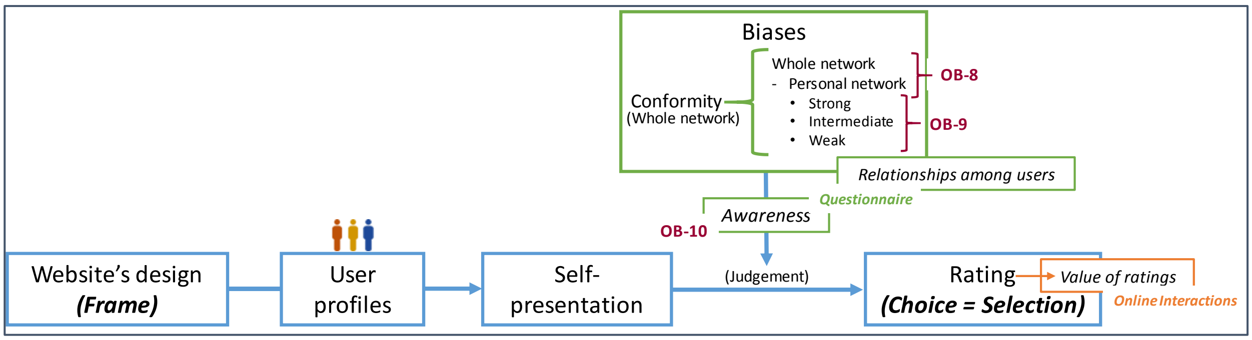


Figure 6.1 - Objectives of the third research question and link with the conceptual model

6.2 METHOD

6.2.1 Participants

As in the previous chapter, the present one includes participants from Groups 2 and 3. That is, it comprises a total of 678 third-year undergraduate students, 351 females and 327 males. Group 2 has 369 students, 330 based in the UK and 39 in China; while Group 3 has 309 pupils, 278 based in the UK and 31 in China. In addition, the ratings of Group 1 are also used in the analysis. Given that the first cohort lacked of declared friendships, the ratings among online 'followers' are used to predict personal networks. Group 1 involved 295 users, 267 based in the UK and 28 in China.

6.2.2 Quasi-experimental set-up and data collection

The quasi-experimental set-up was the same as described in Chapter 5. In both Group 2 and 3 participants were all signed-up with their real identities. Further, both cohorts were asked to complete a survey that included, among other questions, to indicate at least three classmates they personally knew, from either within or outside the class. Finally, three focus groups were conducted per group.

6.2.3 Procedure

The procedure for data analysis and description of the variables is as follows:

A. *Determining if participants conformed differently to their personal networks (Figure 6.1, OB-8)*

It can be complicated to ask people how close they think they are to other members of their social group. Therefore, it is common for SNA to ask questions such as 'how often do you see this person?' or 'to whom do you go when you need advice?' (Jackson 2008). Therefore, to classify the ties among participants, these were asked during the questionnaire to name at least three people from their class whom they personally knew, and then to indicate how often they saw each other, by selecting one of the following options:

- *Frequently:* Classmates that you regularly see outside your class, at least once a week; and if you use social media, they are within your social network.
- *Occasionally:* Classmates that you sporadically see outside your class, but you have met them outside the classroom at least once in the last year and/or you know updates from their lives through social media.
- *Only in class:* Classmates you only know from your class, but you have never seen them outside the class and you do not have them in social media.
- *N/A:* If you prefer not to say, simply choose "N/A"

Participants who saw each other frequently were considered to have *strong* ties, and they can be described as close friends. Moreover, those who saw each other at least once in the last year or have one another in a SNS (e.g. Facebook), can be thought to have *intermediate* ties and can be defined as casual friends. Further, those students who only met in class and were not friends in any SNSs could be regarded as having *weak* ties, and they can be considered acquaintances. In addition, if marked with 'N/A', participants were still considered 'friends', but no strength was assigned to their relationship. Finally, if there was no declared connection between pupils, these were considered to have *absent* ties.

It should be noted that it is not common for SNA surveys to determine the strength of ties based on people having one another in a SNS. However, this element was included in the classification of relationships because students are third-year undergraduates, i.e. they are millennials in their early twenties and therefore having someone as a 'Facebook friend' and getting regular updates on each other lives can be quite meaningful.

B. *Exploring different methods within social network analysis (Figure 6.1, OB-8)*

The manner in which SNA is performed starts with a simple analysis, which then increases in complexity. First, the analysis begins by comparing the average ratings among non-friends (i.e. absent ties) and friends (i.e. personal networks), building up from the three-way chi-square from the previous chapter (see Figure 5.8). Second, the strengths of ties are included, so the average ratings among strong, intermediate, weak, and absent ties are compared. Third, making use of the connections declared by participants, a clustering analysis is performed with the use of the software *Gephi*. Clustering helps to unravel not only the dynamics between two users but among groups of people. Hence, this analysis reveals if relationships affect ratings at the group level. Fourth, geodesic distances (i.e. degrees of separation) are obtained for all users included in the network of friendships. Geodesic distances are similar to clusters but, unlike them, they do not allocate people in one group or the other, but merely count the number of paths between any two individuals within a network. Fifth and final, a multilevel model (MM) is obtained. However, although preliminary results are significant, the model is not pursued further because a more complex MM type is required to gain higher accuracy.

It should be noted that clustering, geodesics and MM were firstly applied only to students from Group 2, and only the 'best' of these methods – geodesics – was then applied to Groups 1 and 3. Namely, the precise methods and terminology that were used are as follows:

- *Networks and strength of ties*

As explained in Chapter 3, networks have been defined as "a set of actors connected by a set of ties" (Borgatti & Foster 2003, p.992). For this research, the actors, or *nodes*, are online users. Further, ties connect pairs of actors and can be *directed* (e.g. Twitter followers, which do not need to be reciprocal) or *undirected* (e.g. Facebook friends, where the friendship needs to be mutual), and *dichotomous* (e.g. present or absent, as with Twitter follower) or *valued* (e.g. measured regarding the strength of friendship, as in Facebook's smart lists) (Borgatti & Foster 2003). In the case of this research, the data that was collected through the survey involved directed, valued connections. Nevertheless, for the analysis, ties were assumed to be undirected. That is, if a student A declared that she saw student B frequently, it was assumed that B would also see A frequently unless specified differently by student B. Moreover, once undirected ties were obtained, the analysis was as follows. First, the rating-averages of friends (i.e. undirected, dichotomous ties) and non-friends (i.e. absent ties) were compared. Second, the rating-averages among different valued ties were contrasted. Hence, rating-averages were compared between: absent (non-friends), weak (acquaintances), intermediate (casual friends), and strong ties (close friends). This comparison was performed with the use of histograms and with an ANOVA test.

- *Clustering*

Cluster analysis regards the study of group structure in multivariate data (Dean 2016). For this research, clusters were created using the software Gephi (Bastian et al. 2009). The goal of cluster analysis is to group multivariate data and, unlike classification, it is used when there is no *a priori* group information (Dean 2016). Namely, data is meant to be unsorted and the aim is to investigate whether there are any groups in the data and, if so, how many and what they look like. Hence, there is no 'right' number of clusters (Dean 2016). In order to detect groups, Gephi (2017) uses the *Louvain method*, which is a type of hierarchical clustering that aims to detect communities within a network (Blondel et al. 2008). In order to perform the clustering, all nodes (i.e. students) and links (i.e. undirected friendships) were put into Gephi. This software produced several SNA metrics, but only two are presented in this analysis. 1) The *node's total degree*, which "is the number of links [i.e. total undirected friendships] that involve that node" (Jackson 2008, p.29), and 2) the *community-detection* *algorithm*, which was randomised in order to obtain different combination of clusters. Afterwards, the average ratings were aggregated at the group level, using the different cluster combinations. Further, the ratings nested into clusters were used to 'summarise' (in a quantitative way, with the use of a table) the overall ratings among different groups. Finally, a few demographics were included to explain the interactions between clusters further. It should be noted that, in a further step, these same ratings nested into clusters were used to explore a basic multilevel model that was used to determine, statistically, if in future analyses it would be significant to keep analysing the data nested into clusters.

- *Geodesic distances*

Geodesics have been defined as the shortest paths between two nodes (Jackson 2008). These distances were obtained because, although the clustering analysis provided significant benefits, it has the shortcoming of allocating users in only one group. That is, if a given person (A) is directly connected with two individuals belonging to different groups, A will be 'allocated' with either of them. Conversely, with geodesic distances, there will be the same distance (e.g. 1) from A to each of the other individuals. Geodesics were considered to be undirected, assuming there is the same distance from A to B than from B to A. Moreover, a starting distance of '1' was given to strong and intermediate ties (close and casual friends), whereas initial distances of '2' were given to weak ties and N/As (acquaintances and not defined relationships). This was done to take into account the strength of ties. Once geodesics were obtained, by making use of all the declared friendships and MATLAB, the rating averages were aggregated within the different distances to compare if, on average, the degrees of separation among participants affected the ratings. It should be mentioned that some students had 'infinite distance' among them, meaning that no path could connect them. This happened with certain individuals that did not declare any friendships and no one mentioned them as friends, or when a group of students were only friends among themselves and were not connected to the rest of the network (i.e. to the 'giant component'). Hence, infinite distances were not considered in the analysis. Finally, it should be highlighted that geodesic distances were also used to 'predict' the effect of relationships with the data from Group 1. This was done by following the same procedure described above but, instead of using the declared friendships from the surveys conducted in Groups 2 and 3, the relationships were inferred by using PeerWise's 'follow function' which allowed users to follow other authors (see Appendix 3.1).

- *Multilevel modelling*

The last used method was MM. This method can be used to analyse clustered or grouped data (Buxton 2008). A basic MM was conducted to investigate if there was enough statistical evidence suggesting that data was indeed clustered into groups (Stride 2016). This analysis was performed by aggregating the ratings within the clusters obtained with Gephi. The data was put into SPSS, and the *Interclass Correlation Coefficient (ICC-1)* was obtained to assess the extent to which the total variance could be attributed to between-group differences (Field 2013; Stride 2016).

C. Comparing real-life friendships with online followers (Figure 6.1, OB-9)

As mentioned, geodesic distances were first used for Groups 2 and 3 by making use of the declared friendships collected in the questionnaires. However, as there were no surveys in Group 1, personal networks were inferred by making use of PeerWise's 'follow function'. In order to make these predictions more accurate, not all 'followers' were assumed to be friends. Instead, the following procedure was used to infer undirected friendships among any two users of PeerWise:

1. *At least one* participant should be following the other (e.g. undirected).
2. Also*, both participants* should have *either* the same location *or* the same nationality[29]. This condition assures that students at least know each other from class; meaning that, as a minimum, they have weak ties and therefore are acquaintances.
    - There were very few cases in which 'followers' were from different locations *and* nationalities. However, even though there was a possibility that they were friends due to exchange programs, these 'followers' were not considered to predict friendships because there was also a chance that they were 'following' each other either randomly or due to the content of their questions.

[29] Different location but same nationalities were allowed because some Chinese students, as part of their programme, had to take to take some modules in Sheffield and some in Shanghai. This meant that they had at some point taken classes together in either of the locations. Indeed, all of the students that were following each other but were not in the same location had Chinese nationality.

D. *Understanding the extent to which online users are aware of the impact that others have on their choices (Figure 4.1, OB-10)*

In order to understand the extent to which students were aware of the impact that their personal networks had on their choices, data from the questionnaire and focus groups were used. Hence, this section presents histograms and quotes from students.

6.3 RESULTS

As with the previous two chapters, the structure of the present results section is built around the outlined objectives. First, conformity to personal networks is tested through the statistical comparison of the outgroup (i.e. absent ties) versus the ingroup (i.e. personal network); and further by introducing and analysing the effect of the strength of ties: strong, intermediate, weak, and absent. Second, personal networks are studied in-depth through the use of SNA and community detection algorithms. This section explores and compares different methods: clustering analysis, geodesic distances, and multilevel modelling. Third, after comparing these methods, the most adequate (i.e. geodesics) is used to compare declared friendships with PeerWise's 'followers'. Hence, followers are first compared with friendships for Groups 2 and 3, to detect if they are somehow similar. Thereafter, once followers proved to be an accurate predictor of real-life friendships, the 'follow function' is used for the data of Group 1, in an attempt to understand if friendships had any effect on ratings when pupils were signed-in anonymously. Finally, the last section seeks to explain the extent to which users are aware of the effect that their personal networks have on them.

6.3.1 Conformity to Personal Networks

<u>Summary of the collected data for mapping the friendship networks of Groups 2 and 3</u>

Table 6.1 presents the number of declared friendships in Groups 2 and 3. As can be appreciated, the questionnaires from Group 3 helped to map a more comprehensive network – including 99.4% of students – than in Group 2, where only 84.3% of students were comprised. Further, it should be noted that only ratings among students are being considered, i.e. questions authored by the "Admin" have been taken out of the analysis. The reasons for this are, firstly, that questions posted by the module's teaching team had already been analysed separately because they involved some sort of 'real-life prestige'. Secondly, the staff of the module could not have been a 'friend', as they did not have personal relationships with the students.

Table 6.1 - Declared personal networks of Groups 2 and 3

Declared friendships	**Group 2**	**Group 3**
Students active in PeerWise ['nodes']	369	309
Directed (declared) friendships ['links']	756	1,270
Bidirectional friendships (duplicated)	192 (96 pairs)	456 (228 pairs)
Undirected friendships (non-duplicated)	1,416	2,312
Percentage of the class comprised in the friendship's network	311 students = 84.3%	307 students = 99.4%
Total number of ratings between students	25,491 (100%)	14,686 (100%)
Ratings btwn non-friends [30]	23,443 (91.97%)	12,723 (86.63%)
Ratings btwn friends (personal network)	2,048 (8.03%)	1,963 (13.37%)
Average rating between all students	2.94	4.69
Ave. rating btwn non-friends (absent ties)	2.88	4.66
Ave. rating btwn friends (personal network)	3.61	4.89

<u>Conformity towards personal networks</u>

When comparing conformity between absent ties and personal networks, two points should be highlighted regarding the table above. First, in both cohorts, the number of ratings among 'non-friends' (i.e. absent ties) were greater than the number of ratings among 'friends' (i.e. personal networks). Second, also in both groups, the average ratings given between 'friends' were higher than for 'non-friends'. These findings suggest that pupils were mostly accessing

[30] It should be reminded that 'non-friends' comprise absent ties, whereas 'friends' (i.e. personal networks) encompass strong, intermediate, and weak ties; plus 'N/A' (non-specified friendships, see Section 6.2.3). The following section makes a further breakdown of the strength of ties.

content that was not created by their ‘friends’. However, when they did come across content authored by someone within their personal network, they would rate it higher.

Quantitatively, the difference between the ratings given to outgroups in comparison to those from within pupils’ ingroups was very significant. In the case of Group 2, which used a likert scale to rate questions, the mean rating among friends ($M=3.61$, $SE=.020$) was significantly higher than that between non-friends ($M=2.88$, $SE=.006$). The mean difference ($\Delta= -.733$) was very significant at $t(2394.5)= -35.75$, $p<.001$. Likewise, in Group 3 – where the dichotomous rating scale was used – friends were rated with an average of 4.89, while non-friends received 4.66. The difference in rating averages ($\Delta= -.223$) was concluded to be significant, with the use of an independent-samples Mann-Whitney non-parametric test: $U = 13{,}044{,}465$, $z = 7.65$, $p < .001$.

Moreover, there was also qualitative evidence in questionnaires and focus groups that showed that the closeness to other participants influenced ratings, as well as other online interactions such as comments. For instance:

> *“I found that my friends would generally rate the question highly even if it was a poor question, and the people I didn't know would rate the question harshly. However I think everyone found this to be the same, and I guess it is the natural thing for people to do”* (sic., Quest-G2, cyfr444).

Likewise:

> *“if you saw a username and it was one of your friends’, you'd think about what you'd say more carefully. Whereas if it was someone you didn't know, it was kind of them being anonymous almost, so you might as well just say what you thought of it”* (FG3-G3, anse680).

Personal networks and the strength of ties

As was shown, both with qualitative and quantitative evidence, students tend to favour their personal networks at the moment of rating questions within PeerWise. Also, as previously outlined, the theory of strength of ties proposes that different 'levels of friendship' translate to different dynamics among individuals (e.g. Granovetter 1973; 1983). Hence, it is worth going a step further and investigate if different levels of relationships affect conformity, by making a further differentiation among personal networks with the strength of ties. Table 6.2 presents a summary of the collected data, regarding the strength of ties, for Groups 2 and 3:

Table 6.2 – Declared strengths of ties for Groups 2 and 3

	Group 2	Group 3
Directed (declared) friendships ['links']	756	1,270
Strong:	281	335
Intermediate:	229	364
Weak:	174	374
N/A:	72	224
Undirected friendships (non-duplicated)	1,416	2,312
Strong:	498	561
Intermediate:	436	659
Weak:	342	654
N/A:	140	438

Then, the strength of the relationship between 'author' and 'rater' was obtained for every rating done in PeerWise, using the undirected friendships shown in the table above. Table 6.3 shows the aggregated average ratings per strength of tie:

Table 6.3 – Ratings per strength of tie: count, percentage, and average

Friendships	**Group 2**			**Group 3**		
	No. of ratings	**Perc.**	**Ave. rating**	**No. of ratings**	**Perc.**	**Ave. rating**
All ratings among students	25,491	100%	2.94	14,686	100%	4.69
Non-friends (i.e. Absent ties)	23,929	94%	2.88	13,034	89%	4.66
Friends (personal networks)	2,048	8%	3.61	1,963	13%	4.89
Strong ties	1,177	5%	3.73	1,014	7%	4.94
Intermediate ties	561	2%	3.59	579	4%	4.83
Weak ties	264	1%	3.32	300	2%	4.83
N/A	46	0%	2.63	70	0%	4.79

The most remarkable findings are, first, that although Table 6.2 shows almost the same number of declared friendships per strength of ties, Table 6.3 shows that the number of ratings was proportional to the strength of the relationship between users. That is, in both cohorts there were more ratings between strong ties, followed by intermediate, weak, and N/A's. This finding suggests that students were more aware – and thus more likely to access – the questions posted by their close friends, as opposed to those authored by their acquaintances. However, if ratings between absent ties are considered, these were by far the highest percentage in both groups; indicating that students were in no way only tackling and rating questions posted by their personal networks.

Second and most importantly, for both cohorts, the stronger the tie, the higher the rating. This finding can be observed both in Table 6.3, above, and Figure 6.2 shown below. Also, for both Group 2 and 3, the strength of ties turned out to be very significant. In Group 2, the ANOVA of the average values for the strength of ties gave the following results: $F(3) = 485.36$, $p<.001$, and an estimated R^2 value of 5.4%. Likewise, for Group 3 the test showed $F(3) = 20.95$, with an R^2 value of 0.4%. However, it should not be forgotten that the results between the ties of Group 3 are not as easily differentiated as those from Group 2 – both visually and statistically – because of the dichotomous rating scale. Note that the complete set of results can be found in Appendix 6.1

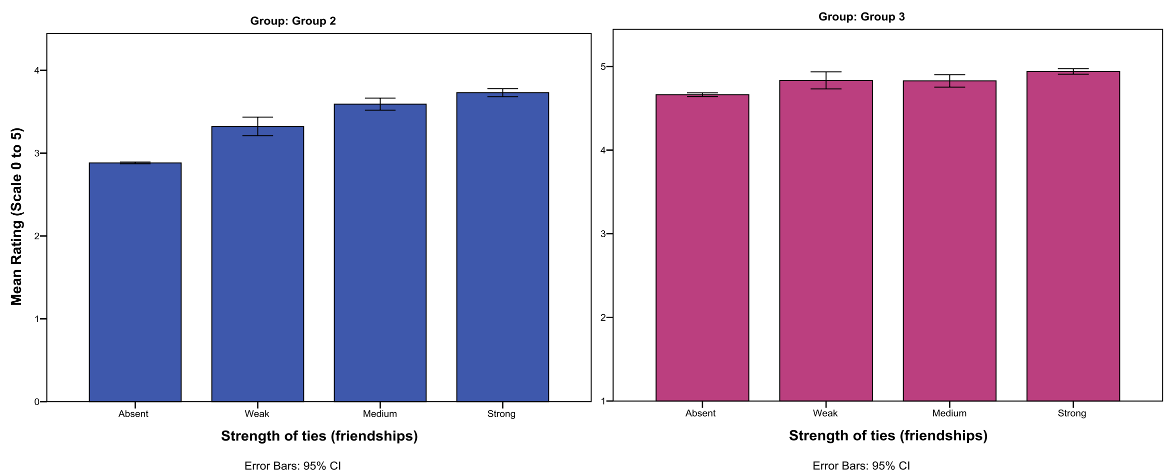


Figure 6.2 – Histograms of average rating per tie strength, for Groups 2 and 3

6.3.2 Personal Networks: Exploring Communities

The previous tables and figures already demonstrate that there was a significant difference between the ratings of 'friends' and 'non-friends', and also between the strengths of ties. However, Table 6.1 and Table 6.3 also show that the percentage of ratings between declared 'friends' was rather small in comparison to the totality of ratings. Moreover, the R^2 values shown above reflect that the strength of the ties can explain little of the variation in ratings. As mentioned, this could mean that pupils were not aware of their friends' questions – although they would rate them higher when encountered –, or perhaps it can be attributed to the shortness of the rating scales, especially in Group 3. However, another reason could be the lack of data. For instance, the response rate of the questionnaire was 50% in Group 2 and 60% in Group 3, meaning that not all students declared friendships. Further, students were asked to name at least three 'friends' within the class, but it is likely that they had more. Consequently, the declared friendships are thought to be fewer than in real-life. For this and other reasons, it was decided to detect groups or communities within the network through the use of SNA.

It should be noted again that communities were first obtained only for Group 2 using a range of methods, such as clustering, geodesics, and multilevel modelling. Afterwards, when the 'best' method was selected (i.e. geodesics), this was applied to the declared friendships of Group 3 and was then used to predict relationships in Group 1. The reasons for using the data from Group 2 as a basis were, firstly, that it was available one year before that of Group 3. Secondly, the Group 2 had a broader rating scale which made small differences more noticeable. And thirdly, the second cohort had fewer declared friendships than Group 3 – both in total and in proportion –, which meant that its analysis could receive a greater benefit from detecting communities, as these may uncover undeclared relationships.

Clustering (Group 2)

As described in Section 6.2, students (i.e. 'nodes') and their declared friendships ('links') were input to Gephi to obtain clusters through Louvain's method (Blondel et al. 2008), which is a type of hierarchical clustering. It should again be mentioned that there is no unique way to determine an 'optimal' number of clusters (Dean 2016). Therefore, different combinations were obtained by randomising the modularity function in Gephi (i.e. its community-detection algorithm). Appendix 6.2 shows an image of all nodes and links from Group 2 before the clustering analysis was performed. This image shows the 82% of students that were connected to the rest of the group (i.e. the network's 'giant component') plus the 58 students who had no declared friendships on the surrounding of this giant component. Further, Figure 6.3 presents only the giant component, differentiated by cluster and showing the total degree of nodes.

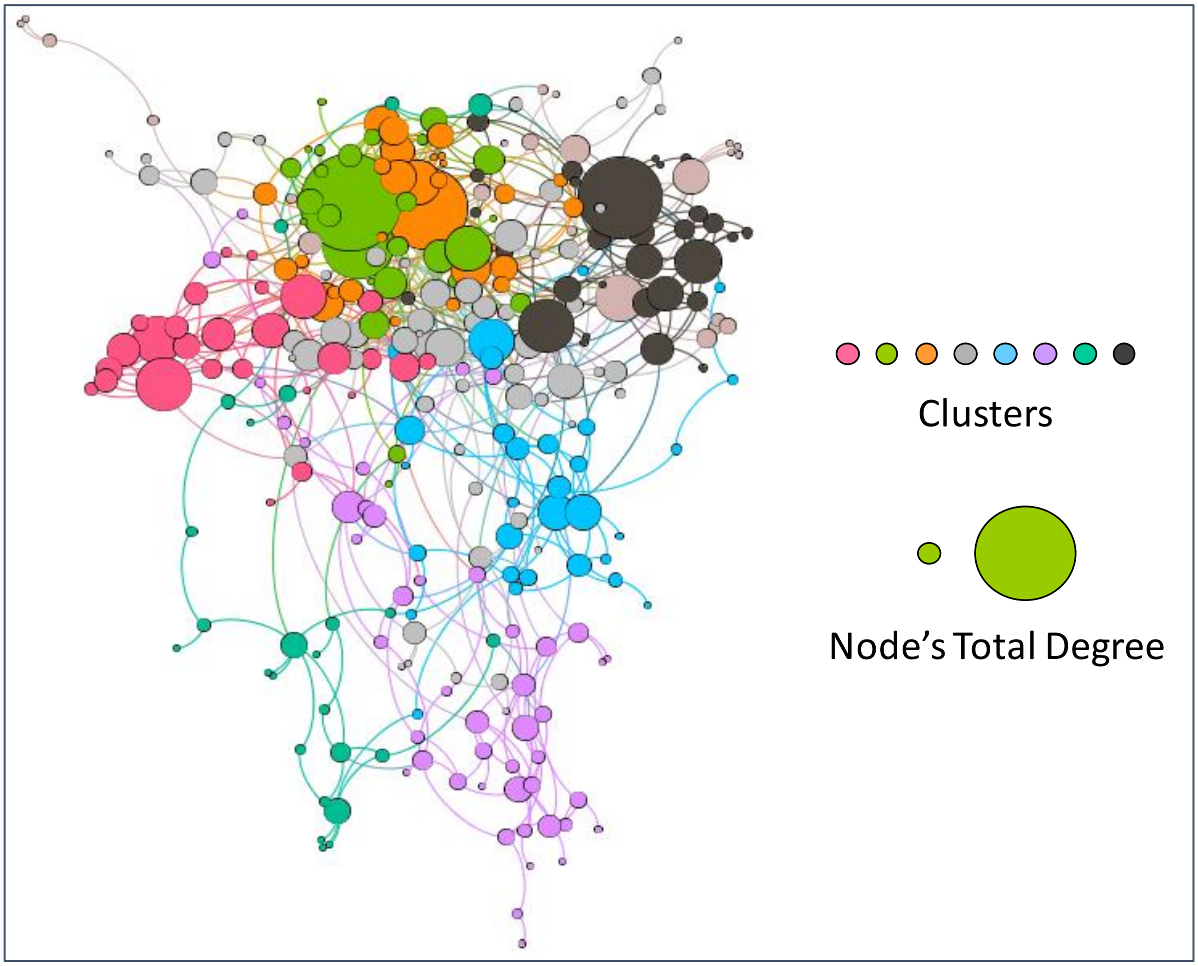


Figure 6.3 – Giant component of Group 2, showing differentiated clusters and network centrality

Furthermore, regarding cluster analysis, the minimum number of clusters that could be obtained was 4: one giant component (see above), two small clusters that were not connected to the rest of the network, and a fourth 'group' which included the 58 pupils with no connections. Conversely, the maximum number of clusters was 78, which comprised 20 clusters of nodes with at least one connection, and an ‘individual cluster’ for the 58 nodes without links. Moreover, with the randomised modularity described in Section 6.2, six cluster arrangements were attained and analysed: 4, 10, 12, 15, 21, and 78 clusters. The most remarkable finding was that, even with the smallest arrangement (4), which produced the largest communities ranging between 28 and 109 pupils each, every single cluster rated itself higher than the rest of the group. This was true for every arrangement. As an example, Table 6.4 shows the rating averages between clusters. The arrangement of 12 clusters was chosen to be displayed because it is a middle-point among the attained cluster arrangements.

Table 6.4 - Rating averages between clusters (Group 2, 12 clusters)

Rater/Author	Author Unassigned	Author Cluster 1	Author Cluster 2	Author Cluster 3	Author Cluster 4	Author Cluster 5	Author Cluster 6	Author Cluster 7	Author Cluster 8	Author Cluster 9	Author Cluster 10	Author Cluster 11	Author Cluster 12	Grand Total
Rater Unassigned	↑ 3.00	↓ 2.71	↓ 2.86	↓ 2.80	↑ 3.01	↓ 2.73	↑ 3.10	↓ 2.71	↓ 2.88	↑ 2.97	↑ 3.03	↓ 2.41	↓ 2.82	↓ ***2.90***
Rater Cluster 1	↓ 2.60	↑ 3.27	↑ 2.96	↑ 3.04	↓ 2.76	↓ 2.92	↓ 2.77	↓ 2.65	↑ 3.01	↓ 2.77	↓ 2.87	↑ 3.17	↓ 2.69	↓ ***2.87***
Rater Cluster 2	↓ 2.79	↑ 3.02	↑ 4.12	↑ 3.09	↓ 2.83	↑ 3.07	↓ 2.79	↓ 2.70	↓ 2.84	↓ 2.58	↑ 3.08	↓ 2.50	↓ 2.64	↑ ***3.01***
Rater Cluster 3	↓ 2.78	↑ 2.95	↑ 3.01	↑ 3.78	↑ 2.98	↑ 3.02	↓ 2.76	↓ 2.61	↑ 3.08	↓ 2.63	↑ 3.60	↓ 2.33	↓ 2.44	↑ ***3.08***
Rater Cluster 4	→ 2.93	↓ 2.70	↓ 2.91	↓ 2.87	↑ 3.41	↓ 2.87	↓ 2.56	↓ 2.63	↓ 2.74	↓ 2.54	↑ 3.12	↓ 2.80	↓ 2.44	↓ ***2.84***
Rater Cluster 5	↓ 2.72	↓ 2.82	↑ 2.97	↓ 2.86	↓ 2.61	↑ 3.52	↓ 2.43	↓ 2.52	↑ 3.06	↓ 2.47	↑ 2.97	↓ 2.44	↓ 2.24	↓ ***2.85***
Rater Cluster 6	↑ 3.10	→ 2.93	↑ 3.05	↑ 3.01	↓ 2.91	↓ 2.86	↑ 3.33	↓ 2.72	↓ 2.86	↓ 2.81	↑ 2.98	↓ 2.46	↑ 3.15	↑ ***3.04***
Rater Cluster 7	↓ 2.58	↓ 2.70	↓ 2.86	↓ 2.66	↓ 2.70	↓ 2.56	↓ 2.54	↑ 3.14	↓ 2.73	↓ 2.36	↓ 2.73	↓ 2.75	↓ 2.23	↓ ***2.73***
Rater Cluster 8	↓ 2.60	↑ 2.97	↓ 2.81	↑ 2.99	↓ 2.59	↑ 3.05	↓ 2.61	↓ 2.65	↑ 3.29	↓ 2.63	↑ 3.14	↓ 2.11	↓ 2.44	↓ ***2.84***
Rater Cluster 9	↑ 2.98	↑ 3.13	→ 2.93	↓ 2.92	↓ 2.84	↓ 2.90	↓ 2.92	↓ 2.77	↑ 2.99	↑ 3.48	↑ 2.94	↓ 2.33	↓ 2.89	↑ ***3.08***
Rater Cluster 10	↓ 2.76	→ 2.94	↑ 2.96	↑ 3.21	↓ 2.88	↑ 3.07	↓ 2.90	↓ 2.63	↓ 2.62	↓ 2.69	↑ 3.54	↓ 2.50	↓ 2.83	↑ ***2.98***
Rater Cluster 11	↓ 2.88	↓ 2.71	↑ 3.00	↓ 2.67	↑ 3.25	↑ 3.00	↓ 2.33	↑ 3.50	↑ 3.00	↑ 3.29	↓ 2.50	↑ 4.20	↑ 3.00	↑ ***3.06***
Rater Cluster 12	↓ 2.79	↓ 2.80	↑ 3.00	↓ 2.80	↓ 2.72	↓ 2.72	↑ 2.95	↓ 2.67	↓ 2.77	↓ 2.86	↑ 3.05	↓ 2.17	↑ 3.37	↓ ***2.88***
Grand Total	↓ ***2.87***	↓ ***2.91***	↑ ***3.07***	↑ ***3.05***	↓ ***2.90***	↑ ***2.95***	↑ ***2.96***	↓ ***2.77***	↑ ***2.94***	↑ ***3.00***	↑ ***3.12***	↓ ***2.58***	↓ ***2.85***	→ ***2.93***

Table 6.4 presents the average rating given by the 'raters' of a certain cluster to the 'authors' of that and other clusters. The cells highlighted in yellow show the highest value per row; that is, they highlight the highest rating average that each cluster gave. As mentioned above, all clusters rated the questions authored by their members higher than the rest of the class. The only exceptions were the 'unassigned' pupils, who rated those in Cluster 6 the highest. Moreover, the arrows pointing upwards show when a cluster rated another one above average; below average if these were pointing downward; and at the average if they were horizontal. From this, it was discovered that clusters which were more similar to one another, regarding location and nationality, rated each other higher. Conversely, when students from other clusters were not similar to themselves, ratings were below the average. This similarity by nationality is indicated by the colour of the headings of rows and columns: red if the cluster had a majority of Chinese students, blue if these were mainly from the UK, and green if they belonged to other nationalities. This finding could suggest that, after friendships, students gave higher ratings to those who were more similar to them. However, this was not true for every single cluster. For instance, Cluster 5 rated those in Clusters 1 and 3 below the average, although the majority of students in them had the same location and nationality. Therefore, this finding rather suggests ratings were mainly based on friendships (and not homogeneity), and nationality was a predictor for friendship.

Geodesics (Group 2)

Clusters were very helpful in detecting communities and identifying how these behaved in terms of rating one another. However, they presented two main disadvantages. First, as previously mentioned, there is no 'correct' number of clusters, so any findings obtained can always be challenged, and are not easily replicated. Second, and more importantly, although people rarely belong to solely one community, cluster analysis 'forces' them to belong to a single one. This could also explain why, sometimes, similar clusters rated each other above the average, while in other cases they rated each other below the average: people having multiple memberships bring two or more clusters together. For these reasons, it was decided to work with other approaches, such as geodesic distances, which involve counting paths between any two individuals.

Table 6.5 presents the geodesic distances for participants in Group 2. It should be noted that distances that contained a count of 100 ratings or less were not considered in the analysis because their averages generated higher variances. Tables containing the omitted values can be found in Appendix 6.3. Moreover, there are two aspects from the table below that should be highlighted. First, it is worth noting that the average ratings are inversely proportional to the geodesic distance among participants. Namely, the fewer degrees of separation between two people, the higher the ratings – at least for distances 1 to 6, where it stabilises until a distance of 10. These results were significant at $F(9) = 254.55$, $p<.001$. Second and most important, in comparison with the strength of the ties where only 8% of ratings were among personal networks (see Table 6.3), geodesics comprise 76% of ratings. Hence, it might have a higher explanatory power.

Table 6.5 – Geodesic distances between pupils of Group 2

G2 – Geodesics from Friendships			
Distance	**No. Ratings**	**Pct.**	**Ave. Rating**
GeoD_1	1,743	7%	3.68
GeoD_2	1,938	8%	3.38
GeoD_3	2,504	10%	2.99
GeoD_4	3,471	14%	2.87
GeoD_5	3,646	14%	2.77
GeoD_6	2,977	12%	2.76
GeoD_7	1,765	7%	2.79
GeoD_8	883	3%	2.79
GeoD_9	311	1%	2.92
GeoD_10	117	0%	2.77
Sum	*19,355*	*76%*	
All ratings, students	**25,491**	**100.0%**	**2.94**

Comparison of methods

Table 6.6 presents a comparison of the used methods in Group 2, conducted at the question level. The most suitable method are geodesics because, as shown in the table below, these comprise the majority of ratings and hence account for a higher percentage of the variation in ratings. Moreover, geodesic distances present the advantage that their results are always consistent and replicable; that is, the shortest path between two people is always the same. Moreover, the same person can have as many connections as the number of members in the network; implying that nodes are not allocated to one or the other cluster, but instead have a fixed geodesic distance with every other node within the network. The full results of each regression, including the unstandardized coefficients, can be found in Appendix 6.4.

Table 6.6 – Regression for SNA methods, Group 2

	No. of Ratings	Percentage of ratings	Coefficient of determination	Unstandardized coefficient	Significance
Prop. of Friendships	2,047	8%	$R^2 = .050$	Beta = .224	t = 36.675 , p<.001
Ave. Strength of Ties	2,047	8%	$R^2 = .054$	Beta = .232	t = 38.116 , p<.001
Prop. of Same Cluster	4,070	16%	$R^2 = .059$	Beta = .243	t = 39.932 , p<.001
Ave.Geodesic Distance	19,355	76%	$R^2 = .069$	Beta = -.262	t = -37.778, p<.001

Multilevel Modelling (Group 2)

Before replicating geodesics for the other cohorts, it is worth mentioning the results of the MM briefly. The intention of performing this analysis was to determine if it would make sense to treat the data as 'nested' within a series of higher-level units (i.e. clusters), and if analysing it in this manner would be a 'better' method than the ones presented above.

MM helps to determine the amount of variance explained at each level of the model (Field 2013; Stride 2016). In the case of Group 2, the first level comprised the ratings given by participants, and the second level involved the clusters where participants were nested. The data used were the arrangement of 12 clusters presented in Table 6.4. Two models were run. First, an 'unconditional model' (i.e. with no predictors), to determine whether a model with varying intercepts was suitable; the unconditional model only partitions the variance in the dependent variable (Stride 2016). This first model produced an ICC-1=0.270, which means that 27% of the variance can be attributed to the group differences, i.e. clusters. Therefore, it can be said that it makes sense to analyse the data in a multilevel manner. Moreover, this would mean that respondents within the same cluster are more alike than respondents from different clusters (Stride 2016).

Second, a model was ran including 'same cluster' as a level-2 predictor. This model showed an improvement from the first one, reducing its -2LL *(-2 Log Likelihood*), hence giving a reduction in deviance of 1,421 on 2df, and showing a very significant chi-square at $p<.001$. Most importantly, it showed that 'same cluster' was an important and significant predictor for 'rating score'. Results for both MMs can be found in Appendix 6.5. However, and despite these significant results, it was not possible to advance in the creation of a MM because the ratings provided a difficulty as they were within and between clusters, repetitive, and bidirectional. Hence, a special type of MM needed to be implemented, which is called 'multiple memberships, multiple classifications' (e.g. Tranmer et al. 2014). However, the software for analysing these types of models (e.g. MLwiN) is still under development and, although it allows for interactions within and between clusters, it does not currently allow for bidirectional, repetitive measures. For this reason, it was decided not to follow the MM path at the present moment. Additionally, the four methods compared in Table 6.6 already fulfilled the posed research question and objectives.

Geodesic distances of Friendships (Groups 2 and 3)

After determining that geodesics were the most suitable method for understanding the impact of friendships on ratings for Group 2, this same method was applied to the data of Group 3. Table 6.8 presents the geodesic distances for participants in Group 3. These results were very significant at $F(6) = 17.64$, $p<.001$. Moreover, as can be seen, geodesics cover 96% of the total count of ratings, a much higher percentage than the 13% covered by the strength of ties shown in Table 6.3.

Table 6.7 – Geodesic distances between pupils of Group 3

G3 – Geodesics from Friendships			
Distance	**No. Ratings**	**Pct.**	**Ave. Rating**
GeoD_1	1,606	11%	4.90
GeoD_2	2,292	16%	4.82
GeoD_3	3,235	22%	4.68
GeoD_4	3,958	27%	4.62
GeoD_5	2,152	15%	4.61
GeoD_6	722	5%	4.57
GeoD_7	150	1%	4.63
Sum	*14,115*	*96%*	
All ratings, students	**14,686**	**100%**	**4.69**

Similar to the results from the geodesics of Group 2, the findings from Table 6.7 show that the rating of questions was inversely proportional to the geodesic distance between two pupils – again, at least among distances 1 to 6. Therefore, it can be deduced that, for both cohorts, personal relationships play a significant role in ratings, and not only among close or casual friends but up to six degrees of separation. Further, having a separation of three degrees or less between author and 'rater' would almost guarantee that the question is rated at-or-above the group's average. Hence, given that the ratings within PeerWise did not count towards students' grades, it can be argued that the closer a person is to oneself and one's personal network, the higher the perceived quality of their information.

6.3.3 Inferring friendships from online 'followers'

As was described in the literature review, a theoretical and practical gap regarding the strengths of ties is that there is a need for joining social media models with real-life data (Gilbert & Karahalios 2009). Researchers commonly use 'followers' from different SNSs or CCs (e.g. Twitter) to infer real-life relationships, but rarely do they get the chance to verify if the predicted relationships are meaningful.

Concerning this research, the previous section demonstrated how geodesic distances provide a solid basis for unravelling personal networks within groups, and the effect these have on the ratings of participants. However, given that not all students answered the questionnaire, these could have made the geodesics obtained by friendships somehow incomplete. Also, Group 1 did not complete a questionnaire to capture personal networks, and it has so far been assumed that these had no impact on the ratings because students were anonymous. However, it was also discovered that when personal factors were introduced to the regression, nationality turned out to have a small but significant impact on the ratings of the first cohort (see Table 4.7 and Table 4.8). For all these reasons, 'followers' are used to predict friendships.

Following peers: opinions from questionnaires and focus groups

As was previously described at the start of this chapter, PeerWise's 'follow function' is used when a user wants to 'subscribe' to the questions authored by a specific individual (see Appendix 3.1). Thereafter, when the 'followed' peer authors a new question, this is highlighted to the 'follower'. Moreover, given that 'following' peers was a way in which students could select questions, they were asked about this functionality during the questionnaire and focus groups. In both Group 2 and 3, this function was ranked sixth (out of eight) as a method of choosing a question to answer. Hence, based on this, it would seem that it was not among the most popular methods for targeting questions.

Subsequently, students were asked in the surveys if they had ever used the follow function and, if they did, whether they had followed based on content (i.e. quality of questions) or friendships. The results were as follows. In Group 2, 80.3% of students (N=183) stated they had used the 'follow' function, and from these, 60.1% said that they had mostly followed people they personally knew. Likewise, in Group 3, 74.6% of students (N=205) claimed to have followed authors within the website. From these, 56.9% confirmed that they mostly followed students they knew in person, rather than because of the quality of the author's questions. Namely, in both groups about half of the students (48.3% in Group 2 and 42.4% in Group 3) declared to have followed someone from their personal network. This discovery was also confirmed during the focus groups, where some students were even blunter about their use of the 'follow' function: *"I was doing [the assignment] for the sake of doing it. I just used [PeerWise] 'cause I had to. And I would answer my friend's questions only. I'd follow them, we'd follow each other and just do us, so we didn't really engage"* (FG1-G3, elya294).

Comparing friendships and followers (Groups 2 and 3)

The personal networks obtained with the questionnaires were compared with those obtained with PeerWise's 'followers' and the results were somehow similar. Namely, in Group 2, 29.8% of the undirected 'friends' were also undirected 'followers'; and this percentage was 21.1% for Group 3. Hence, about a third of the students who declared a friendship were also following one another in PeerWise.

Moreover, the ratings of PeerWise's 'followers' were compared with the ones of 'non-followers'. The results looked similar to the comparison between 'friends' and 'non-friends' presented at the start of this chapter, in Table 6.1. Also, what appeared rather surprising was that the 'followers' seemed to have a similar structure and effect on ratings than that of the strength of ties (see in Table 6.3). This finding can be observed in Table 6.8 below, which was obtained by dividing 'followers' into 3 categories: those who were both following one another (bidirectional), those where only the 'rater' was following the 'author' (directed), and those where the student was being followed and was targeting the followers' question (indirect).

Table 6.8 – Number of collected ratings and followers for Groups 2 and 3

Followers	**Group 2**			**Group 3**		
	No. of ratings	**Perc.**	**Ave. rating**	**No. of ratings**	**Perc.**	**Ave. rating**
All ratings among students	25,491	100%	2.94	14,686	100%	4.69
Non-followers	19,954	78%	2.74	10,682	73%	4.61
Followers	5,537	22%	3.67	4,004	27%	4.92
Bidirectional	3,155	12%	3.84	2,033	14%	4.96
Directed	1,812	7%	3.47	1,607	11%	4.89
Indirect	570	2%	3.38	364	2%	4.88

As can be observed, similar to the strength of ties, the different 'levels of followers' present proportional rating averages. Moreover, the percentage of ratings among 'followers' exceeds those attained with declared friendships, when only 8% and 13% of ratings were among personal networks, respectively. In contrast, as can be seen in the table above, the ratings between followers were 22% for Group 2 and 27% for Group 3.

Inferring friendships from followers (Groups 2 and 3)

With the criteria described in the method section of this chapter, geodesics are obtained for 'followers' and then compared the geodesics obtained from declared friendships (see Table 6.5 and Table 6.7), in order to investigate if personal networks could be predicted by online 'followers'. Group 2 comprised 1,243 'followers', from which 1,212 had either same location or nationality. Group 3 had 969 'followers', of which 919 were from the same location or same nationality. Hence, the average rating per geodesic distance of the 'inferred friendships' predicted using online 'followers' are as follows:

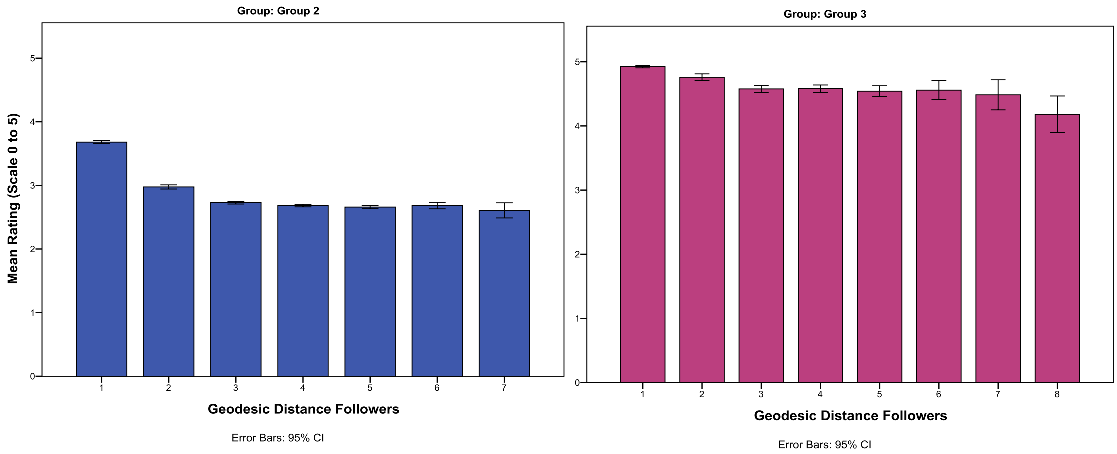


Figure 6.4 – Rating averages per geodesic distance for 'followers' in Groups 2 and 3

Similar to the geodesic distances of friendships, the ones for 'followers' turned out to be very significant. For Group 2, the impact of follower's geodesic distance was significant at $F(6) = 1,020.78$, $p<.001$. Likewise, Group 3's distances were significant at $F(7) = 37.18$, $p<.001$. The full results can be found in Appendix 6.6. What is more, as more of the ratings were given between 'followers' than among 'friends' (i.e. a greater percentage of ratings were comprised by 'followers' than by personal networks), the explanatory power of geodesics obtained by 'followers' is higher than those attained by declared friendships. Using as an example the second cohort, if geodesic distances of 'friends' were put into a regression model, these would explain 6.9% of the variation on ratings. In contrast, geodesic distances of 'followers' can explain 15.4% of ratings. Hence, given that Group 1 had no declared friendships and because it had the same rating scale than Group 2, the geodesic distances of 'followers' could be used as a predictor of the influence that personal networks had on ratings, if there was one.

Inferring friendships from followers (Group 1)

As has been mentioned, it came as a surprise that nationalities were playing a small yet significant role in the ratings of Group 1 (see Table 4.8). Hence, at the time of analysing the online interactions, it was speculated that some students could have worked in teams outside the class or used other technology, like WhatsApp groups, to send links of their questions to friends. Still, this could not be proved as there was no chance to know the students' opinions through questionnaires or focus groups. Yet, there was one student in Group 2 who commented on the role of personal relationships in Group 1 (see below).

> *"...because I did a placement the majority of my friends did this [module] last year and they were saying how they worked together, commented on each other, and agreed to answer [their questions]. But I don't know anyone in this course anymore, I've got like three mates and it would be too obvious almost if I answered all of my friends' questions. Because then there would be only three of us just bouncing questions around off each other. And it's not fair because, what if you had no mates on this course? Some people could have chosen this as an elective and had no friends and they would be at a disadvantage because they would not be able to create this 'team effect'"* (FG3-G2, ieje546).

For this reason, it was considered worthwhile to explore how the friendships in Group 1 might have influenced ratings, and how this influence compared to that of the other two groups. The same criteria to infer relationships from Groups 2 and 3 were used for the first cohort. Group 1 had 1,214 'followers', from which 1,155 shared either the same location or nationality. Geodesic distances were obtained and, from the 295 active students in the group, 273 (92%) were comprised in the network. Figure 6.5 shows the average ratings per geodesic distance.

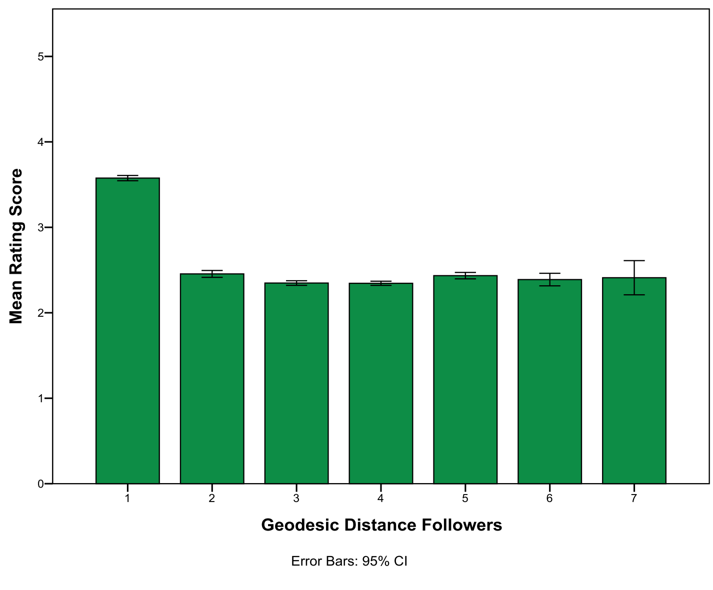


Figure 6.5 – Rating averages by follower's geodesic distances for Group 1

As can be seen in Figure 6.5, ‘followers’ (i.e. ‘inferred friendships’) did play a role in Group 1, but only at the first level; that is, only at a distance of one. When all distances are included in the test, ANOVA proves very significant at $F(6) = 935.56$, $p<.001$, and $R^2 = 11.4\%$. However, if the distance of 1 is removed from the analysis, the ANOVA becomes non-significant, and the R^2 value goes to 0.0%. Suggesting that, if friendships had an effect on the ratings of Group 1, this only happened between close, direct friends (i.e. friends with one degree of separation). However, the questions of friends-of-friends (e.g. two degrees of separation and more) did not receive higher ratings, unlike the other two cohorts. Table 6.9 shows the inferred ‘friendships’, obtained with the ‘follow’ function, for all the three groups.

Table 6.9 – Geodesic distances between followers of Groups 1, 2, and 3

Followers	**Group 1**			**Group 2**			**Group 3**		
Distance	**No. R**	**Pct.**	**Ave. R**	**No. R**	**Pct.**	**Ave. R**	**No. R**	**Pct.**	**Ave. R**
GeoD_1	5,741	23%	3.58	5,466	21%	3.68	3,847	26%	4.92
GeoD_2	2,662	10%	2.45	2,700	11%	2.98	1,598	11%	4.76
GeoD_3	5,325	21%	2.35	5,070	20%	2.73	2,423	16%	4.58
GeoD_4	6,447	25%	2.35	5,466	21%	2.68	2,270	15%	4.58
GeoD_5	2,941	12%	2.43	2,948	12%	2.66	1,135	8%	4.54
GeoD_6	741	3%	2.39	893	4%	2.68	362	2%	4.56
GeoD_7	117	0%	2.41	201	1%	2.61	165	1%	4.48
GeoD_8							165	1%	4.18
Sum	*23,974*	*94%*		*22,744*	*89%*		*11,965*	*81%*	
All ratings, students	**25,490**	**100%**	**2.64**	**25,491**	**100%**	**2.94**	**14,686**	**100%**	**4.69**

There are a number of findings that deserve to be commented. Firstly, for Groups 2 and 3 it is quite surprising that even inferred relationships show the same pattern as real friendships: geodesic distances are inversely proportional to ratings, and significantly different up to distances of 6 or more. Secondly, it is interesting that this did not happen in Group 1, where there is only a significant difference between distances of 1 and all the others, but not among distances of 2 to 7. This suggests that, because the website was set-up anonymously but still had the ‘follow function’, students took care in finding who their friends were, and ‘followed’ them. However, as the rest of the group were anonymous, weaker ties (i.e., longer geodesic distances) did not play a role on their ratings. This can also be confirmed by the fact that in this cohort only ratings coming from the first geodesic distance were rated above the average, whereas for Groups 2 and 3 this happened for distances of 1 and 2.

This last section provides evidence to support the argument that personal networks influenced the ratings of students from the three cohorts. Moreover, it also reveals that anonymous users were only influenced by their ‘close friends’. Conversely, identifiable users seemed to be rating almost out of friendship, inversely proportional to the degrees of separation from their peers. The next and final section assesses the extent to which students were aware of the influence that others exerted on them.

6.3.4 Awareness of conformity to personal networks

In Group 2 (N=183) students were asked to select all aspects which they considered most pupils had taken into account at the time of rating a question. Results were as follows:

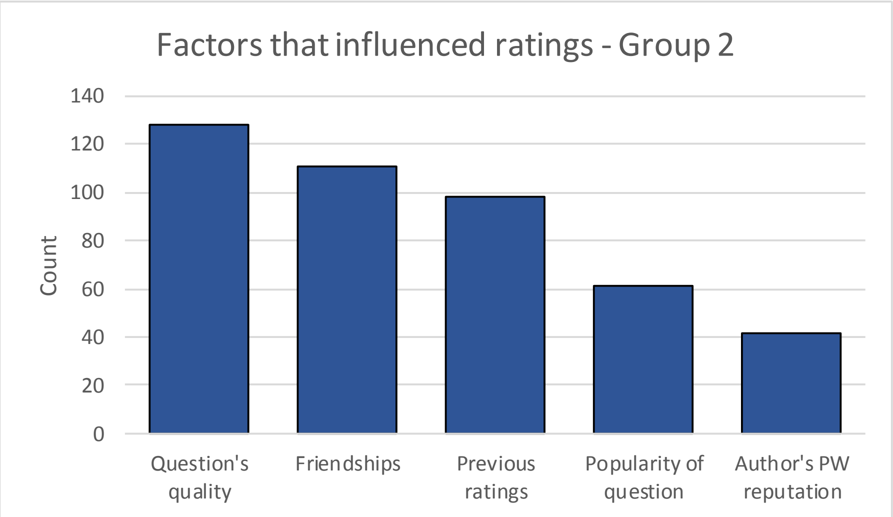
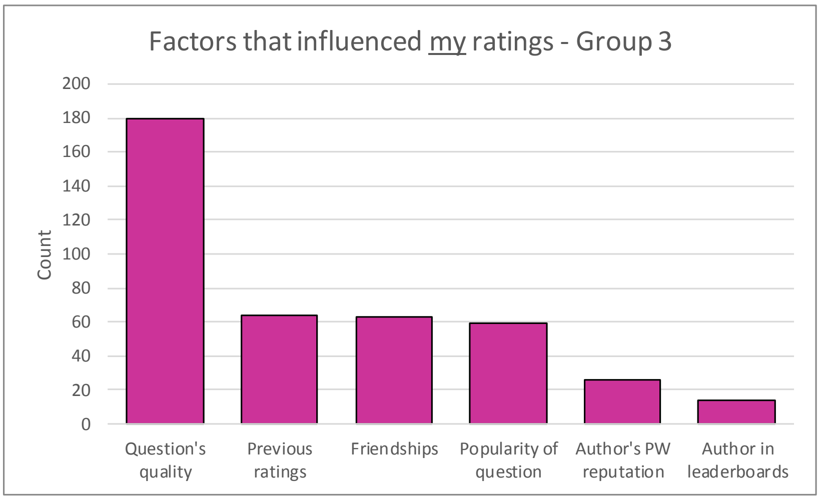


Figure 6.6 – Factors that influenced the ratings of ‘others’, Group 2

A refinement to the questionnaire of Group 3 (N=205) separated the above question into two: the first asking about the respondent and the second about other users. Figure 6.7 shows the disparity in answers. As can be seen in, participants believed (or at least reported) that they had been fairer and had based their ratings mostly on quality, with their personal networks having only a small effect, and ranked in third place behind previous ratings of a question. However, to their eyes, their peers gave less importance to the quality of questions, and relied almost equally on friendships, followed by previous ratings, popularity and reputation scores.



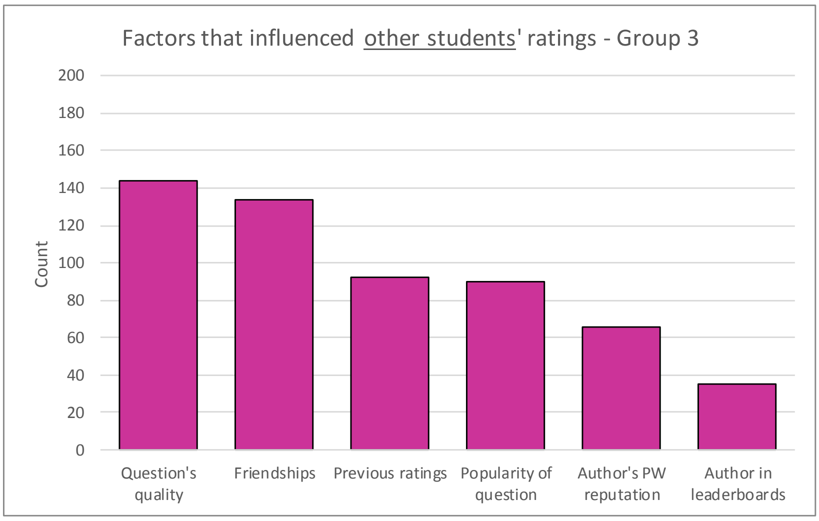


Figure 6.7 - Factors that influenced the ratings of oneself VS others, for Group 3

Likewise, participants' comments on both surveys reflect what the previous figures show. In the questionnaires, many students raised the issue of friends helping each other, but in all the cases they referred to 'others' or 'everyone', but rarely to themselves: *"People I noticed were liking and giving high ratings to their friends questions – which ultimately makes the legitimacy of the rating score unreliable"* (Quest-G3, iego907). The only cases where students agreed helping their friends, was when they were asked to do so: "*My friends expected that i would rate their questions as excellent but i am not like them because i want to be fair. And it was very difficult for me to decide whether i'd "lie" or not on Peerwise and rate something that i dont believe it's true*" (Quest-G2, nist698).

Conversely, during the focus groups, students were more open and some were very sincere about 'playing the system', notwithstanding that their module marks did not depend on this: *"You could, like, subscribe to particular users so a group of my mates* *did subscribe because we thought we could help each other out, like, almost cheating the system. So, we were like 'oh, I just posted a question, can you answer it and write a comment I could reply to?"* (FG2-G2, ntbu009).

Nevertheless, although almost all students accepted to have mainly answered their friends' questions and/or rated their friends higher, they honestly added that their friends' questions were of a good quality. For instance, (from FG1-G2):

- *Researcher: I actually wanted to ask about that. What would you do if you saw one of your friends' questions and you thought 'this is (a bad question)'. Would you not answer it or would you let them know it's bad?*
- *lyro673: "This might sound funny, but the questions I answered from my friends were actually quite good. So I was like 'nice one, this was good'. I didn't see any bad questions… I guess I would have probably not answered it. Because they had lots of questions, so I would have just gone to a different question. You know? No one would get annoyed if you didn't answer one question".*

In conclusion, the findings from this section suggest that students were, to some extent, aware of the influence that diverse factors had on their ratings. Yet, as Figure 6.7 shows, they believed that their rating choices had been mostly based on the quality of the questions, while other aspects such as friendships, had only played a minimal role. However, it was easier for them to detect the biases that influenced their peers, and they were quite aware of these. Even the most honest students, who openly 'bragged' about 'playing the system', added that their friends' questions were actually of decent quality. However, the analyses from the online interactions presented in all of the results chapters have shown how everyone was affected by different biases, the strongest one being the effect that personal networks had on them.

6.4 SUMMARY OF FINDINGS

The third research question dealt with conformity, towards the whole and personal networks. Its purpose was to explore if – and how – this takes place. Moreover, this section also examined whether online users were aware of the influence that others had on them, and to what extent. This chapter outlined three objectives. The first to be addressed (OB-8) was to determine if – and how – users' real-life ingroups and outgroups affect the way in which they perceive and evaluate UGC. As mentioned, ingroups consisted of users' personal networks (i.e. strong, medium, and weak ties), whereas outgroups comprised absent ties between any two users. Similar to previous studies (e.g. Dobson et al. 2013), it was found that the transmission of knowledge was favoured by stronger and closer connections. Specifically, it was found that in Groups 2 and 3 the strength of the tie was directly proportional to the count and mean-value of the ratings. Therefore, between weak, intermediate, and strong ties; the latter one accounted for the largest number of ratings among students, in addition to having the highest scores, on average. However, some relationships were unaccounted for because not all students completed the questionnaire, and those who did were only required to declare a minimum of three friendships. Hence, it was decided to use clustering and geodesic distances in order to unravel relationships that had perhaps not been declared.

Thereafter, the ratings of students were firstly arranged within clusters, and it was found that every cluster rated its members with the highest scores, followed by clusters with individuals of similar locations or nationalities, and lastly the remaining clusters were given ratings below average. However, although clustering analysis provided helpful insights, it presented the issue of allocating participants in only one cluster, regardless of the connections that the user might have had with individuals from distinct communities. Then, to overcome this issue, geodesic distances were obtained with declared friendships, and it was found that the distances between individuals were inversely proportional to the ratings. Moreover, the average ratings were significantly different for distances ranging from one to six degrees of separation, for both cohorts. Secondly, geodesics showed that those who were separated by a distance between one and three were rated at-or-above the average of the group. Thirdly, in comparison to the other methods, geodesic distances comprised the majority of the ratings of each cohort, making them the SNA variable with the highest explanatory power. Therefore, geodesics were found to be a good predictor which can easily be replicable but, most importantly, they served to elucidate on the effect that different degrees of personal networks had on the perceived quality of online content (i.e. posted questions).

The second tackled objective (OB-9) was to test if online interactions could predict real-life relationships. To achieve this, PeerWise's 'follow' function was used due to participants disclosing to have ‘followed’ their friends to facilitate the access to their questions. By making use of the declared friendships and the list of followers from Groups 2 and 3, it was discovered that at least a third of students used this function to ‘follow’ their friends. Moreover, when comparing the geodesic distances obtained from friendships to those obtained by followers, the similarities were striking. Hence, given that ‘followers’ seemed to be a good predictor for real friendships, the geodesics from ‘followers’ were also obtained for Group 1, which would make possible to explore if ‘friends’ had an impact when participants were anonymous.

The most relevant finding was to compare the results among the three groups. In the first cohort, there was a significant difference between ratings from those with one degree of separation and the rest of the group. Conversely, in Groups 2 and 3 each additional degree had a lower average rating score. Moreover, in Group 1 only those ‘followers’ comprised in the first geodesic distance were rated above the average, whereas in the other two cohorts this happened for distances of magnitude one and two. Therefore, these findings suggested that, possibly, participants who were signed-in anonymously might still have used the ‘follow’ function to differentiate between the content from those who are close to them and the rest of the group. However, as opposed to Groups 2 and 3 – where participants were identifiable –, weaker ties and friends-of-friends did not affect the ratings of anonymous users.

The last objective (OB-10) was to understand if users were aware of the influence that their personal networks exerted on their choices. While it was found that students were aware of the influence of friendships to a certain extent, they showed a much higher awareness of their peers’ biases than their own. From their perspective, the element they took into account at the moment of rating was the quality of the question, followed by a minimal influence from other factors. Conversely, the ratings of ‘others’ were mostly affected by friendships, previous ratings, the popularity of the question, and the author's online prestige. Similarly, in the questionnaires and focus groups, students often complained about their classmates being biased by some of these elements. Specifically, participants seemed especially annoyed when their peers up-voted their friendship groups. Even so, the online interactions showed that all participants were prone to these biases, not just a few of them.

CHAPTER 7: THE EFFECT OF THE WHOLE NETWORK ON INDIVIDUALS' CHOICES, IN THE CONTEXT OF DIFFERENT WEBSITE DESIGNS

As of today, if an individual wanted to go to London for a weekend, (s)he would be presented with the following options. To rent an apartment, she would need to choose from 1,393 family households, 1,115 work houses or 188 'plus homes' in Airbnb. Otherwise, in order to book a hotel, she would need to select one of the 2,563 offered in booking.com. Moreover, when choosing where to eat and what to do, she would need to select one of 18,373 restaurants; followed by one of the 1,265 tours or 1,787 'top things to do' listed in TripAdvisor. Online data is doubling in size every two years and "by 2020 the digital universe – the data we create and copy annually – will reach 44 zettabytes, or 44 trillion gigabytes" (Turner et al. 2014, p.1). As a consequence, online consumers are seeking heuristic information cues to simplify the amount of data involved in taking decisions (Park & Nicolau 2015).

Three reasons were given in Chapter 2 as to why people need these information cues. First, from an individuals' perspective, there is a limited amount of information that people can process in order to make fast, simple, and unconscious decisions that have allowed the species to survive (Kahneman 2012). Namely, the human brain is "wired to avoid complexity (not embrace it) and to respond quickly to ensure survival (not explore numerous options). In other words, our evolved decision heuristics have certain limitations" (Bonabeau 2009, p.45). Second, from a social perspective, humans have evolved allowing individuals to forego the costs of individual learning, enabling people to acquire knowledge through social learning processes like teaching, language and imitation (Mesoudi 2011). Third, the number of available possibilities that exist nowadays overwhelm and paralyse people (Schwartz 2009).

In addition, trusting the content in online environments has become an issue, especially for the new generations. Millennials – the generational group with the highest presence in social media – perceive news reports, company websites, and advertising as significantly less trustworthy than the information shared by their personal relationships

(The McCarthy Group 2014). Specifically, millennials tend to search for online ratings and comments before purchasing goods or services, and they place higher confidence in the opinions available on the network than on those that come from the brands and companies selling the products (Gutiérrez-Rubí 2014).

However, when it comes to relying on the network to make choices, researchers have found mixed results. On the one hand, some scholars claim that the 'wisdom of the crowds' can outsmart the single individuals conforming them (Surowiecki 2004). Likewise, it is claimed that the 'collective intelligence' can outwit individuals, drawing from examples like Wikipedia and Google (Malone et al. 2009; Malone et al. 2010). What is more, some researchers argue that UGC, such as reviews and ratings, can be used to obtain extremely accurate results (Hill & Ready-Campbell 2011). On the other hand, it has been alleged that crowds can be counterproductive in terms of decision-making. Even Surowiecki (2004), who proposed that crowds could be smarter than individuals, outlined some circumstances in which the group-intelligence could fail, such as when people imitate each other or when they are persuaded by a 'leader'. Similarly, researchers have highlighted limitations regarding smart crowds, such as homogeneity and lack of verity (Roman 2009), self-confidence (Bonabeau 2009), and social influence (Lorenz et al. 2011).

This chapter aims to explore the extent to which people rely on the whole network to make choices, and whether this changes in the context of different website designs. To achieve this, the data from the questionnaires and focus groups is used to uncover the diverse ways in which participants relied on the network. Moreover, online interactions from the three quasi-experimental conditions are assessed using sequence analysis (SA) to investigate if social influence (i.e. conformity to the group) had an effect on ratings, and this was different depending on each condition. It should be emphasised that a secondary purpose of this chapter is to compare the three quasi-experimental conditions and summarise all findings, to better introduce the Discussion chapter that follows.

7.1 OVERALL RESEARCH QUESTION AND OBJECTIVES

While the previous result chapters have partially addressed the overall research question of the thesis, this final chapter tackles it directly: *How is the transmission of UGC affected by the different designs (i.e. frames) adopted by social media websites?*

Likewise, although Chapters 4 to 6 have already addressed all of the thesis' objectives, the present chapter draws on some of them to make an overall comparison of the three quasi-experimental designs (i.e. frames) of this research (see Figure 3.1 and Figure 3.5). Figure 7.1 shows how these objectives fit with the full conceptual model of the thesis.

- **OB-2:** To understand knowledge transmission in social media through the VSR mechanisms; emphasising selection, which is studied through choice-making.
- **OB-3&5:** To investigate if – and how – different combinations of user profiles *(OB-3)* and rating scales *(OB-5)* affect the choices of users.
- **OB-7:** … [To focus] on the study of conformity, which has received limited attention both offline and online.
- **OB-4&6:** To determine how both tested attitudes – towards being identifiable *(OB-4)* and to each rating scale *(OB-4)* – affect the overall perception of the website's design.

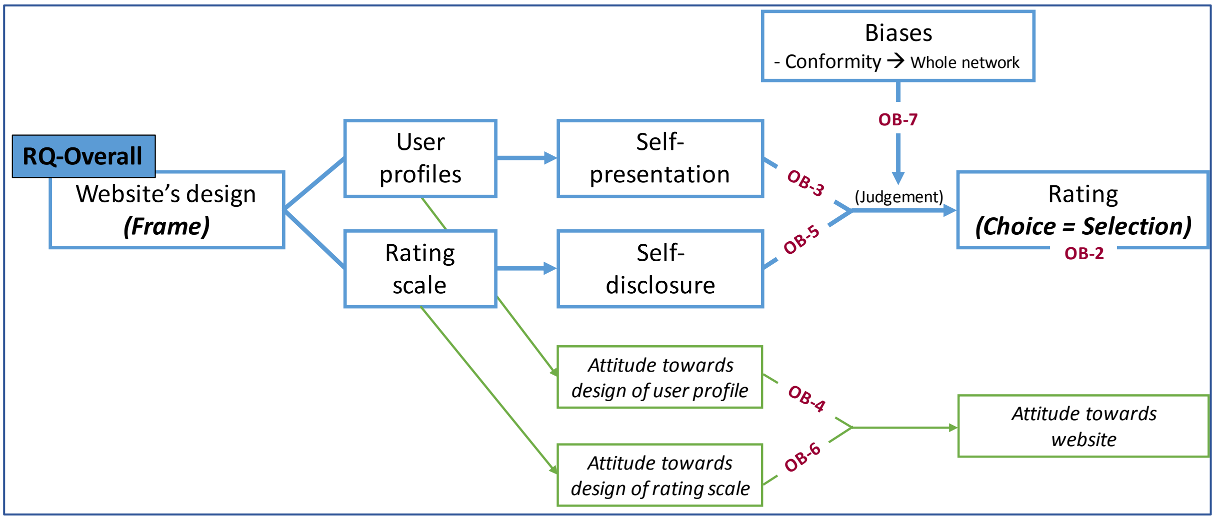


Figure 7.1 - Objectives of the overall research question and link with the conceptual model

7.2 METHOD

7.2.1 Participants

All participants were included in this last chapter. Therefore, the sample includes 973 third-year undergraduate students. These were from 56 different nationalities, although the grand majority were British (48.7%), followed by Chinese (28.1%). The average age was 21.4 years, and 53.1% of the sample were females, while 46.9% males.

7.2.2 Quasi-experimental set-up and data collection

This chapter makes use of the data from three quasi-experimental conditions. As has been explained in Chapters 3 to 6, users in Group 1 were anonymous and could rate on a likert scale. In Group 2, participants were signed-in with their real identities and continued to use a likert scale. Finally, Group 3 was set up so that users were still identifiable but the rating scale changed to a like/dislike dichotomy. Thereafter, when targeting a question, the information available to participants from each cohort was as follows: (Appendix 3.1 shows how the three quasi-experimental conditions affected other aspects of PeerWise).

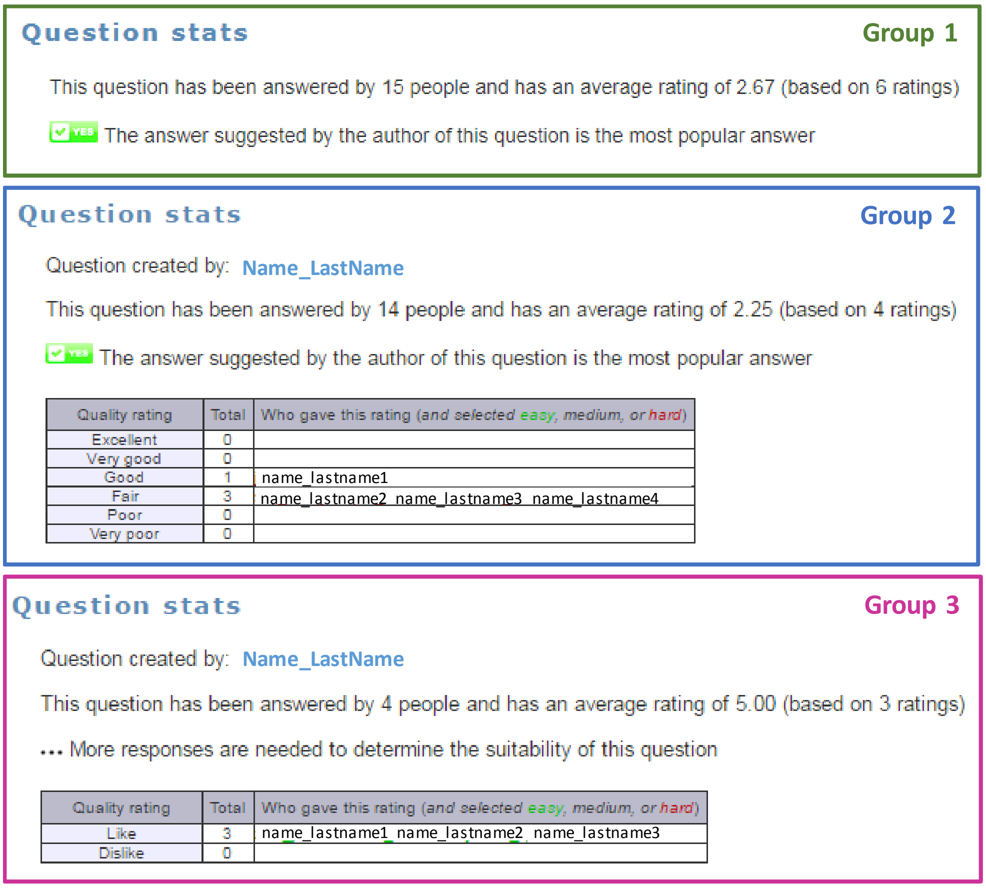


Figure 7.2 – Summary of previous ratings shown to users when targeting a question, per group

7.2.3 Procedure

The procedure for data analysis and description of the variables is as follows:

A. *Understanding knowledge transmission in social media through the VSR mechanisms, making an emphasis on selection (Figure 7.1, OB-2).*

This section makes use of the answers from questionnaires and focus groups to elucidate on the different ways in which pupils relied on the network, and how this impacted the VSR mechanisms. Unlike the other sections, this one does not focus on the ratings *per se*. Instead, it explores other aspects that were prior to ratings, such as authoring questions and selecting which ones to answer. Most of these aspects were impossible to detect through the analysis of online interactions, so the only way to recognise them was through the insights of participants.

B. *Investigating how the different combinations of user profiles and rating scales affect the choices of users (Figure 7.1, OB-3&5).*

The way in which the three quasi-experimental conditions are to be contrasted is through the comparison of the sequences produced with the ratings of each cohort. SA helps to detect patterns in categorical sequences, focusing on the state of the sequence, where "the position of each successive state receives a meaningful interpretation" (Gabadinho, Ritschard, Mueller, et al. 2011, p.1). For this study, a *sequence* is composed of the successive ratings that were given to a particular question. Thus, the *number of sequences* equals the number of authored questions, while the *length* of each sequence is equivalent to the total number of ratings that the question received. Moreover, the different *states* are equivalent to the possible ratings; hence, there were six possible states in Groups 1 and 2 (0 to 5) and two states in Group 3 (0=dislike or 5=like).

The data used to conduct SA were the consecutive ratings, arranged by question, from all three cohorts. Ratings were analysed with *TraMineR* (Gabadinho, Ritschard, Studer, et al. 2011), which is an open-source software that can be installed and run in R. For each sequence (i.e. question), the software produced the measures listed below, which were later aggregated by cohort in order to make a comparison between the three groups. It should be noted that all definitions were obtained from the TraMineR user manual (Gabadinho, Ritschard, Mueller, et al. 2011):

- *Number of transitions:* It is the count of changes of state in a sequence. For example, if two sequences are compared, each of length 5, Seq_1= [1,1,1,1,1] and Seq_2= [1,2,3,4,5], then Seq_1 would have zero transitions while Seq_2 would have four.

- *Within entropy:* Within (or longitudinal) entropy is a normalised measure that describes the diversity of states within a sequence. Hence, when the state remains the same during the whole sequence – e.g. if all the given ratings to a question are the same (like the example of Seq_1) – then the sequence has an entropy equal to zero. Conversely, the maximum entropy is reached when the same proportion of cases are spent in each state. Thus, if each rating within the sequence has a different value (like Seq_2), then the entropy of the sequence is equal to one. This measure does not account for the state order in the sequence.

- *Complexity index:* This measure combines the number of transitions in the sequence with the within entropy. The minimum value of zero can only be reached by a sequence with a single distinct state (i.e. with no transitions and entropy of zero). Moreover, it reaches its maximum of one if, and only if, the sequence is such that it contains each of the possible states, and the same number of cases are spent in each state.

Thereafter, the procedure was as follows. First, the data was input to TraMineR and the measures listed above were obtained for all the sequences. Second, the individual measures were aggregated by cohort. Third, an ANOVA test was performed to assess if the aggregated mean values were significantly different across cohorts. Finally, two visualisations were produced: one comparing all the sequences between groups, and another one contrasting the ten identical questions[31] that were authored and posted by de module staff.

[31] Only the 10 questions posted by the module staff were compared among groups, because the 5 that were authored by students in the first cohort received significantly less ratings in Group 1, making a comparison between groups impossible. Conversely, the 10 questions authored by the module staff were always the first ones posted on PeerWise, and hence stayed the same amount of time in the website, and received a similar number of ratings.

C. *Focusing on the study of conformity, which has been poorly explored online and offline (Figure 7.1, OB-7).*

In order to fulfil this objective, the 'levels of conformity' of each of the three quasi-experimental conditions were obtained and compared. The outputs given by the software TraMineR included a measure that was used to infer the 'level' or 'strength' conformity to the whole network, through the analysis of ratings. This measure was compared between the three cohorts:

- *Transition rates:* These indicate the probability to switch at a given position from state S_i to state S_j. This measure provides an insight of the most frequent state changes observed in the data. Moreover, it gives an evaluation of the stability of each state, which can be observed through the diagonal on the outcome matrix, that represents the transition rates from a state to itself (Gabadinho, Ritschard, Mueller, et al. 2011).

It should be noted that transition rates are based on *Markov chains*, which are "stochastic models describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event" (Oxford Dictionaries 2018). However, in TraMineR "the rates are assumed to be position-independent" (Gabadinho, Ritschard, Mueller, et al. 2011, p.17).

D. *Determining how both tested attitudes – towards being identifiable and to each rating scale – affect the overall perception towards the website's design (Figure 7.1, OB-4&6).*

Chapter 4 tested the effect of the attitude to being identifiable on the overall perception towards the website. Likewise, Chapter 5 investigated if the attitude towards the rating scale had any impact on the overall perception of PeerWise. Thereafter, this section presents a correlation analysis – with the use of Spearman's rho – which allows for a better understanding of the relationship between these three variables. Moreover, the predictive power of both attitudes is obtained by using Cox & Snell pseudo-R^2.

7.3 RESULTS

The results are organised as follows: first, the different ways in which participants relied on the network is explored through the comments that students expressed through the questionnaires and focus groups. Second, the sequences of the three cohorts are compared with the use of SA measures, such as sequence lengths, transitions, entropy and complexity. Third, the extent to which users conformed is inferred through comparing the transition rates obtained for each quasi-experimental condition. Fourth, the attitudes towards being identifiable and to the rating scale are used jointly to understand how they affected the overall perception of the website.

7.3.1 Reliance on the network and its impact on the VSR mechanisms

This section uses the qualitative data obtained from questionnaires and focus groups to understand different ways in which participants relied on the whole network, and how this might have affected the VSR mechanisms. As was proposed in Chapter 2 (see Figure 2.5), at the network-level, the variation of information can be observed through the UGC posted online (e.g. text, images, videos, emoticons). Moreover, the selection can be inferred by the ratings of users (e.g. stars, likes, dislikes, helpful votes). Finally, the overall ranking displayed on the website (e.g. average ratings, count of likes or helpful votes) can be seen as retention. Hence, by observing PeerWise at the network-level: the variation consists of the posted questions, selection comprises the individual ratings, and retention encompass the ‘wisdom’ that is generated by the network, seen through the rating averages. The way in which the VSR mechanisms are thought of taking place on PeerWise is very similar to the representation of StackOverflow, presented on Figure 2.6.

Authoring questions (i.e. variation)

Although participants were not asked about the influence of others on the creation of questions, there were a few comments during the surveys and focus groups that elucidate on different aspects that may have influenced the authoring process. Some of these comments indicate that students relied on the questions posted by their peers to author theirs. In other words, there is some evidence of the variation mechanism being affected by imitation between members of the network:

> *"I thought it was really useful to start with being able to answer questions 'cause initially I didn't want to post a question and not being the right format. Like, a lot of people were using articles and videos and I didn't initially thought of that. So, being able to structure my questions based on what other people had done, I thought was really good"* [sic.] (FG3-G3, myve209).

Further, there was also evidence that the individual ratings (i.e. selection) and averages ratings (i.e. retention) influenced the creation of questions (i.e. variation). Namely, by receiving feedback from other members of the network, individuals learnt which aspects would give them higher ratings and, presumably, this affected the content that they included in future questions:

> *"A couple of the questions I authored included short videos. These seemed to get higher ratings than the questions involving articles. I don't know whether these questions were of higher quality or whether respondents liked the fact that there was visual content and the questions took less time to answer"* [sic.] (Quest-G2, lala535).

Moreover, the comments from students suggest that both being anonymous/identifiable and using a likert/dichotomous rating scale influenced the authoring of questions. In the case of being anonymous or identifiable, some participants argued that PeerWise should have been anonymous because real names affected the feedback they gave to each other: *"You don't want people to know you have put a negative rating/comment on their question"* [sic.] (Quest-G3, enpe347).

Likewise, the quote presented below – which was given in support of the likert scale – shows that some students edited/deleted questions when they received low ratings from their peers. Consequently, the website's choice of rating scale had an impact in the manner in which students were allowed to provide and receive feedback, which affected the variation of content consequently:

"Most questions on PeerWise have a rating of 5 where outstanding question can not easily be distinguished. A rating system of 1 to 5 allows strong question to stand out. At the same time, it reminds authors of sunstandard questions either delete or improve their questions to maintain quality of questions on PeerWise" [sic.] (Quest-G3, eela414).

Choosing questions (i.e. 'pre-selection')

During the questionnaires and focus groups, pupils were asked about the elements that they took into account at the moment of deciding which questions to tackle. In a way, accessing questions can be seen as a 'pre-selection' because – as obvious as it may sound – accessing a question is a necessary step for later rating it. Moreover, as outlined in the introduction of this chapter, people do not read every piece of information available online. Rather, they rely on available heuristic cues to simplify their choices. Hence, users might access a few posts (e.g. reviews of apartments or restaurants) and, from these, they choose where to go.

Similarly, PeerWise has a number of heuristic cues in place, which simplify decisions but might bias the choices of participants. Namely, pupils were able to sort questions by date, tag/topic, author's reputation (badges), popularity (number of answers or comments), difficulty rating, and overall rating. Further, as described in the previous chapter, the online platform also allows users to 'follow' others, updating them every time someone they 'follow' authors a new question. Finally, after answering a question, PeerWise has a button that reads 'submit rating and go to a random question', which allows students to access information randomly (see Appendix 3.1 for the complete screenshots of PeerWise).

Table 7.1 shows how participants ranked the possible manners to access questions. As can be seen – equally in both cohorts – the second, fourth, and fifth most used aspects for choosing questions were based on the 'wisdom' of previous students targeting the question and giving feedback to the network.

Table 7.1 – Ranking of the different options used by participants to tackle questions, Groups 2 and 3

Ranking	Group 2 *(N=183)*	Group 3 *(N=205)*
(Mostly used) 1	Topic/Tag	Randomly
2	Difficulty	Difficulty
3	Randomly	Topic/Tag
4	Overall rating	Overall rating
5	Popularity (i.e. number of previous answers or comments)	Popularity (i.e. number of previous answers or comments)
6	Follow function	Follow function
7	People I knew from class	People I knew from class
(Least used) 8	Author's reputation	Author's reputation

These findings also came up during the focus groups, where most students reported having used previous knowledge contained in the network to make choices regarding which questions to target. For instance, it was common for pupils to say that they only targeted questions with a high count of answers or comments, whereas others declared that they only chose highly rated questions and avoided the ones with low averages. For instance, the following dialogue presents a segment from a conversation in the second cohort (FG2-G2):

- *Researcher: So, how would you choose questions?*
- *All: "Most answered, most recent, by difficulty, by tag/topic, by overall ratings"*
- *iccr997: "I would see the ['Author's answer popular?' column] and I wouldn't answer a question where the author's answer wasn't popular. I don't know, simply 'cause there's more risk for me to get it wrong if not"*

Likewise, the comments from participants in the third cohort also show that they relied on the 'wisdom of the network': *" ... I would just see the most recent [questions] and I'd see if the question had kind of 20 answers and comments, then I'd look into that one because obviously this meant it was a good question... Well, that's what I think it was. And I'd just ignore the really new questions with no answers"* (FG2-G3, nyho990).

Moreover, the manner in which PeerWise was set up for each of the quasi-experiments also affected how questions were selected. For instance, many of the comments from participants highlight how being identifiable made them more likely to access questions from those known to them. Further, regarding the rating scale, previous chapters presented quotes from students ‘complaining’ about the dichotomous scale hindering the differentiation between good and bad quality questions, caused by the disproportionately amount of ‘likes’. Hence, this made users more prone to relying on other knowledge retained in the network, such as the number of answers and comments (see previous quote). To sum up, this section has elucidated on how the information retained in the network affected the choice of questions (i.e. the accessed UGC).

Rating questions (i.e. selection)

This section shows, qualitatively, how the individual ratings (i.e. selection) were affected by the reliance on the knowledge retained in network (i.e. retention). The following two sections – 7.3.2 and 7.3.3 – will do the same but making use of quantitative data and SA.

Similar as with choosing questions, participants revealed to have sometimes based their choices of rating on previous ratings. That is, sometimes they would access a question and, before even reading it, they would have an estimate on how to rate it, thanks to the ‘overall rating’ displayed on PeerWise (see Figure 7.2). For instance: *“When I think about it, I feel I was like kind of influenced by how other people rated it as well. So when I would see that the majority was 'good' or 'very good' I'd think 'ok, I already have an idea of whether the question is good or not'”* (FG2-G2, lala535). This situation was aggravated – as seen in Chapter 4 – by being identifiable, as users were more affected by the previous given by their personal networks. Additionally, the use of the dichotomous rating scale – studied in Chapter 5 – intensified the reliance on previous ratings by only showing two types of previous ratings: ‘like’ and ‘dislikes’.

To conclude, the evidence gathered both from the questionnaires and focus groups of Groups 2 and 3 suggests that students relied on the network to author questions, choose which of their peers’ questions to tackle, and to estimate how to rate the content posted on PeerWise. Paradoxically, the effect that retention had on variation and selection, further affected retention. That is, the effect that the VSR mechanisms had on each other, formed a continued loop. This idea will be further elaborated on the Discussion chapter that follows (see Figure 8.4).

7.3.2 Comparing how different website designs affect choices

As was discovered, there are different manners in which the users of each group imitated each other and relied on the network. Therefore, it could be possible that, with time, some treats or practices became more prominent for some of the cohorts. For instance, from the focus groups, it appeared that participants from Group 2 gave a special value on including videos on their questions, whereas users from Group 3 included more case studies. That is, it is possible that due to the imitation between members, each cohort decided what a 'good question' should contain. Hence, it is likely that ratings also followed a pattern that characterised each cohort.

As mentioned in Section 7.2, SA helps to detect patterns in categorical sequences (Gabadinho, Ritschard, Mueller, et al. 2011). Hence, SA has been applied to the ratings of the three cohorts in order to detect rating patterns. First, it was applied to all ratings; then to the ten identical questions common to the three cohorts, which were authored by the module staff and were posted online at the start of each academic year.

<u>*Sequence analysis of all ratings*</u>

Table 7.2 shows the total number of questions (i.e. sequences) authored in each group, together with the average values described on Section 7.2.3: sequence length, number of transitions in a sequence, within entropy, and complexity of all sequences within the three groups. It should be noted that, to be considered for SA, the minimum length that a sequence should have is one. In other words, questions should have at least one rating[32]. Hence, the minimum length of a sequence was 1, while the maximum 89.

Table 7.2 – Sequence measures for Groups 1, 2, and 3 (with original ratings)

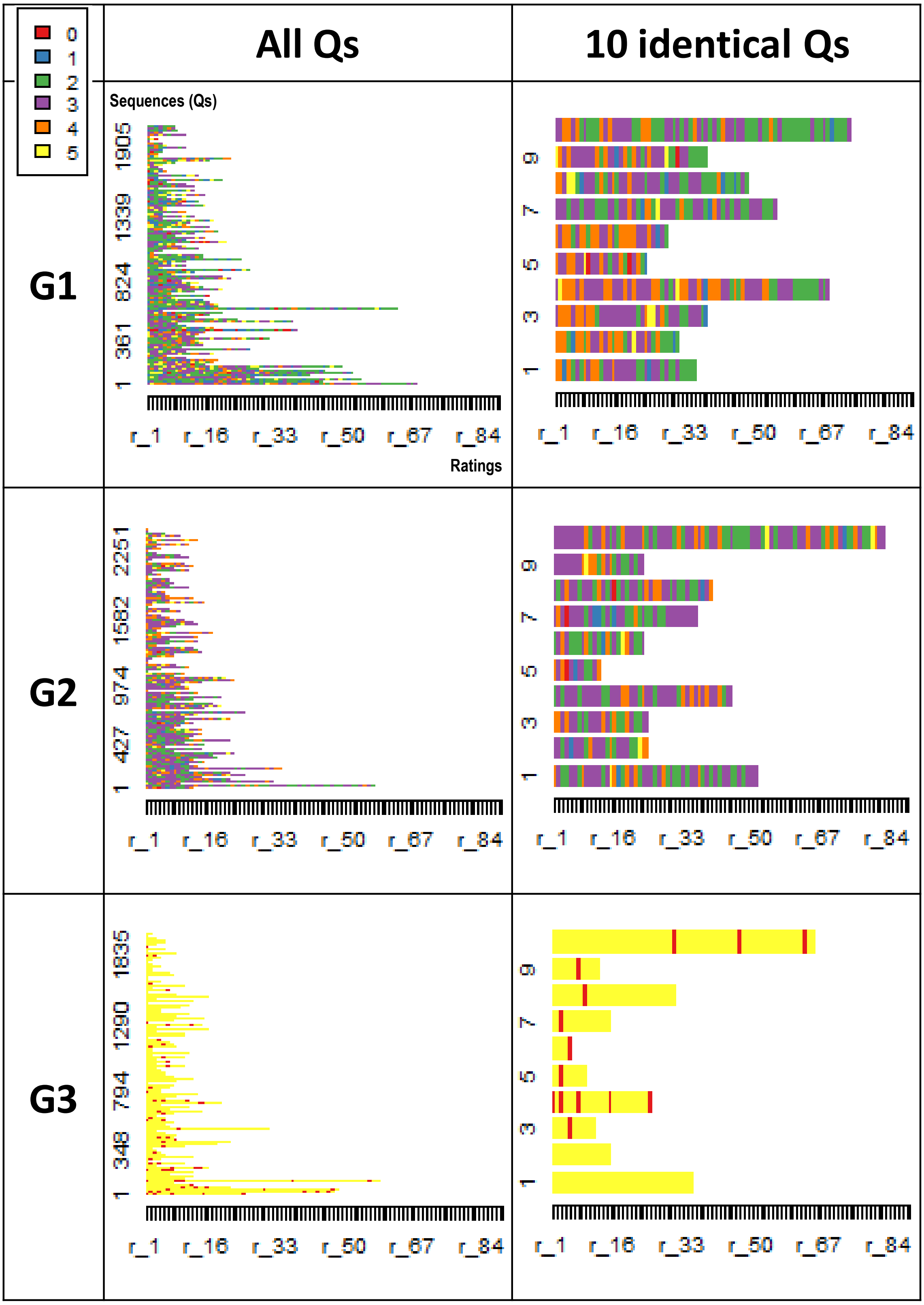
Group	No. Sequences (Questions)	Ave. Length (Ratings)	Ave. Transitions	Ave. Within entropy	Ave. Complexity Index
1	2,064	12.56	7.79	0.59	0.62
2	2,439	10.65	5.77	0.47	0.51
3	1,987	7.55	0.67	0.06	0.07
Total	**6,490**	**10.31**	**4.85**	**0.38**	**0.41**

[32] For this reason, there might be a slight variation between the number of reported questions with the previous chapters.

As can be seen, on average, Group 1 exhibited the highest number of transitions, together with the highest values of within entropy and complexity. This shows that, when students were anonymous and were exposed to a likert rating scale, they tended to change from one rating value to the other with ease. Moreover, identifiable students who used the same rating scale exhibited fewer changes of status, showing higher conformity to the network. Finally, students from Group 3 who were identifiable and had a dichotomous scale displayed the smallest values on each of the three measures, suggesting that they conformed to the network most strongly. An ANOVA test of the differences between groups was performed for each measure, which was highly significant at $p<.001$ in each case. A full set of results can be found in Appendix 7.1.

Moreover, Table 7.3 below shows all the rating sequences by group. In all the graphs, each horizontal line represents a sequence (i.e. question) whose length is measured on the x-axis that comprises all the consecutive states (i.e. ratings). Further, each state (0-5 for Groups 1 and 2, and 0/5 for Group 3) has a colour assigned. Hence, the sequence of colours highlight aspects such as the stability of each sequence and its transitions, which were used to determine the entropy and complexity. Hence, it can be appreciated from the ‘All Qs’ column of the table that the sequences of Group 1 are the longest ones, closely followed by those of Group 2. Sequences of Group 3, instead, look significantly smaller. Moreover, the difference in colours show that Groups 1 and 2 comprise the whole colour palette, but the first cohort looks predominantly green while on the second one the colour that more stands out is purple. Conversely, Group 3 only shows two colours (0 = dislike, in red; and 5 = like, in yellow) but it can still be noticed that there is a predominance of yellow over red, making the disproportion of ‘likes’ and ‘dislikes’ very obvious. Lastly, it is also worth noting the changes in colour, which represent the number of transitions. As can be seen, the cohort in which students were anonymous and used the likert scale has the ‘most colourful’ graph, indicating more transitions between ratings. Then, the cohort where students were signed-in with their real identities and still used the likert scale looks more stable, colourwise. Finally, the third cohort is the one with the longest blocks of one same colour. Arguably, Group 3 looks more stable because only two ratings were used in this graph. However, as explained in the procedure (i.e. Section 7.2.3), the number of transitions is independent of the number of states and only takes into account the number of changes. Nevertheless, to clarify this point, Section 7.3.3 includes another visualisation that converts all cohorts into like/dislike.

Table 7.3 – Rating sequences for Groups 1, 2 and 3 – All & 10 identical questions



Comparison of the ten identical questions authored by the module staff

The sequences from the ten identical questions posted by the module staff show a similar pattern to the one observed for all questions. However, for this visualisation the differences are more noticeable. As can be seen from the '10 identical Qs' column in Table 7.3, the sequences from the third cohort look significantly shorter than the other two groups, despite questions being in PeerWise for the same number of weeks. Moreover, in terms of colour, the sequences from Group 1 start with higher ratings (4, in orange) and then move towards lower ones (2, in green). Further, the ratings of the second group start and remain around the value of 3 (in purple). Conversely, Group 3 comprises mostly 5's (in yellow). Specifically, questions 1 and 2 do not suffer any change in status/rating, which is quite remarkable because it means that not one of the more than 30 students who rated each question 'challenged' its quality, despite not having any personal relationship with the author (i.e. the module staff). Additionally, with the exception of the fourth sequence – which is the only one starting with a 'dislike' rating – users show a tendency to 'dislike' in blocks, rather than sporadically across the question.

These findings, together with those from the previous section, further demonstrate that the way in which websites are designed has an impact on the choices of people. Even when rating the same content, the set-up of the website determines to a significant extent how online users will perceive and select information. By making people identifiable, users probably become more self-conscious and heavily prone to conforming to the network, especially if their friends are present (as shown in Chapter 6). On top of this, when users have few options for rating, they tend only to use those with a positive connotation. Beyond the current educational context, it is worth reiterating that most CCs are making users identifiable. Also, most SNSs – which usually require users to sign-in with their 'real' identities – only allow people to like, love, and favour; and do not even provide an option to show disagreement/dislike or dissent.

7.3.3 Comparing how different website designs affect conformity levels

Although it is not possible to fully determine the degree to which students were influenced by others, SA offers some insight into the level of conformity that students from each group might have experienced. It is worth noting that 'level of conformity' refers to the extent to which pupils were influenced to imitate the ratings from previous 'raters' of a given question. Therefore, differently from Chapter 6 where conformity to personal networks was inferred through the relationship between 'author' and 'rater', the conformity to the whole network is based on previous ratings and hence the analysis is between 'rater' and 'rater'.

Furthermore, although the following analysis does not make a distinction among a rater's ingroup and outgroup, there was qualitative evidence that indicated that both the previous ratings of non-friends and friends affected the ratings of participants. Regarding non-friends: *"…The high reputation and previous rate indeed impact on later rate"* [sic.] (Quest-G2, yuto987). Likewise, concerning the influence of friends: *"…furthermore as previously mentioned friends are more likely to like and dislike same questions"* [sic.] (Quest-G3, ilki790).

The 'levels of conformity' were inferred using the measure of *transition rates,* which is based on the concept of Markov chains, as explained in Section 7.2.3. However, in order to make an even comparison across all three cohorts, the ratings of Groups 1 and 2 were converted to dichotomous similarly to Chapter 5, where the two rating scales were compared. Hence, Table 7.4 shows the basic sequence measures for the three cohorts (as shown in Table 7.2), with all ratings converted to dichotomous. It should be highlighted that two comparisons were made. First, similar to those from Chapter 5, the six values of the likert scale were split in two groups, where 0, 1 and 2 became 'dislike' (D) and 3, 4, and 5 became 'like' (L). This conversion was used to obtain the transition rates. However, it was decided also to present the results for a second conversion where only 0 and 1 were converted to 'dislike' and the remaining four values were equalled to a 'like'. This conversion was used to emphasise the differences between groups through a visualisation shown further below.

Table 7.4 – Sequence measures for Groups 1, 2, and 3 (with converted ratings)

			1st Conversion: [0,1,2]=D and [3,4,5]=L			2nd Conversion: [0,1]=D and [2,3,4,5]=L		
Group	**No. Seq (Qs)**	**Ave. Length (Ratings)**	**Ave. Trans**	**Ave. Entropy**	**Ave. Complex Index**	**Ave. Trans**	**Ave. Entropy**	**Ave. Complex Index**
1	2,064	12.56	4.61	0.74	0.53	2.21	0.39	0.27
2	2,439	10.65	3.42	0.58	0.42	0.60	0.11	0.07
3	1,987	7.55	0.67	0.16	0.12	0.67	0.16	0.12
Total	**6,490**	**10.31**	**2.95**	**0.51**	**0.36**	**1.13**	**0.22**	**0.15**

It is worth noting from the table above that the values for 'number of sequences' and 'average length' are independent of the conversions. Hence, these values are identical to those presented in Table 7.2, whereas the other three measures do vary with the conversions. Regarding the first conversion presented in Table 7.4, for all measures – average transitions, entropy, and complexity – Group 1 shows the highest average values, followed by Group 2 and lastly Group 3. This shows that Group 1 had the most changes between 'like' and 'dislike'. Moreover, concerning the second conversion, Group 1 remained with the highest average values, while Group 2 obtained the smallest ones, albeit not too different from those for Group 3. The results from the second conversion demonstrate that, even when converting the scale in a different manner, Group 1 still presented a significant[33] difference from the other two cohorts. Namely, the effect that website designs have on the ratings (i.e. choices) of its users becomes even more evident. A visualisation of these conversions is presented later, in Table 7.6.

Furthermore, the evidence that each group exhibited a significantly different level of conformity can be reinforced by analysing transition rates (i.e. Markov chains). These were obtained with the first conversion, where half of the likert scale is a 'like' and the other half a 'dislike'. Table 7.5 shows the transition rates for each cohort. As explained, the values show the probability of switching from one state (i.e. rating) to the other within the sequence of ratings, while the diagonals display the stability of each state.

Table 7.5 – Transition rates for Groups 1, 2, and 3

Group 1		
	→ Dislike	→ Like
Dislike →	59.5%	40.5%
Like →	39.2%	60.8%

Group 2		
	→ Dislike	→ Like
Dislike →	40.9%	59.1%
Like →	25.1%	74.9%

Group 3		
	→ Dislike	→ Like
Dislike →	17.9%	82.1%
Like →	5.5%	94.5%

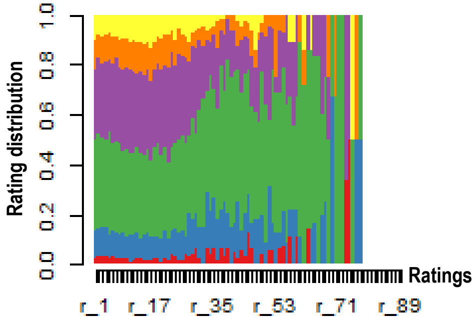
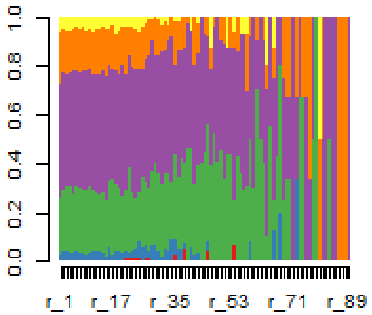
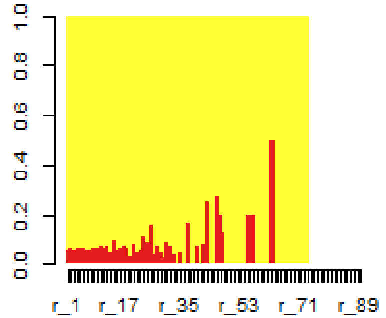
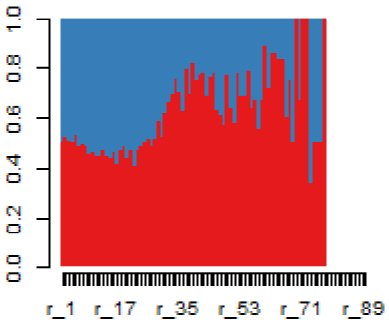
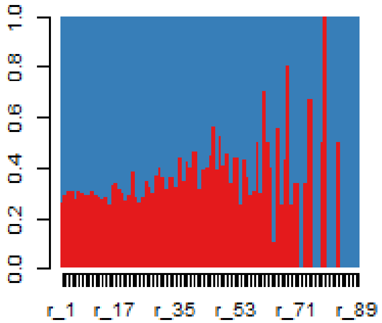
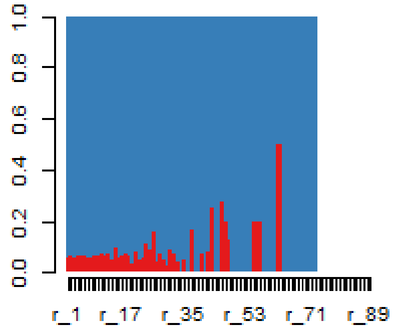
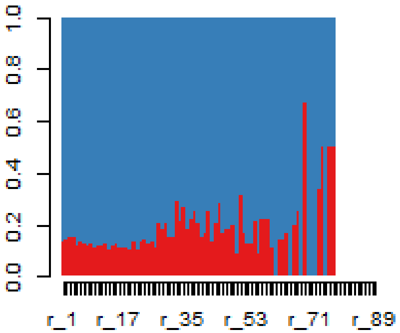
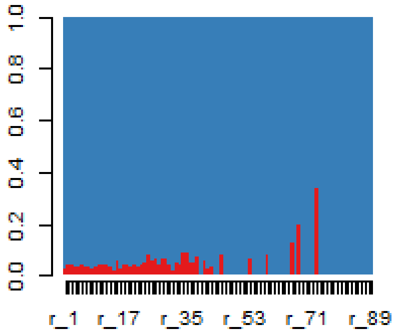
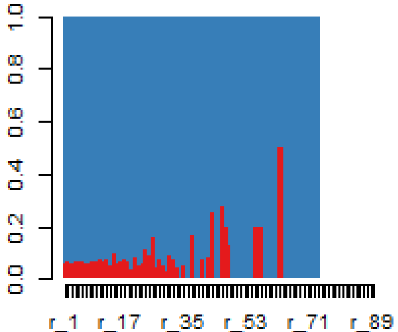
Although the tables above are fairly simple, the insights they provide are invaluable. For instance, they suggest that in Group 1, it was almost as 'easy' to go from one state to the other. That is, there was a 60% chance of continuing with the previous rating value, but also a 40% chance of going from a 'positive' to a 'negative' rating, or *vice versa*. Hence, although there was a slight preference to conform with previous ratings, changing to the opposite value was almost as likely. In contrast, 75% of the users in Group 2 conformed

[33] An ANOVA test of the differences between groups was performed for the measures on each conversion. The results were all significant at p<.001. A full set of results can be found in Appendix 7.1.

to the whole network if the rating was positive ('like'), whereas if it was negative this figure dropped to 41%. Further, going from a 'like' to a 'dislike' was relatively rare, with only 25% of users doing so. Both tendencies are amplified in Group 3, where a 'like' rating was followed by another 'like' 95% of the time, while just 18% of 'dislikes' were followed by another negative rating. Thus, the data shows that conformity to 'positive' ratings (i.e. 'like' → 'like') increases as users become identifiable (in Group 2), and is further intensified by the adoption of a dichotomous rating scheme (Group 3).

To further illustrate the last findings, Table 7.6 compares the rating distributions by group, with three different combinations of the rating scales (i.e. original ratings and the two performed conversions). The distributions represent the percentage of each rating value (i.e. 0-5 and 0/5) per sequence. Furthermore, it is helpful to visualise the effects that anonymous versus identifiable, and likert versus dichotomous scales had on the rating patterns of users. When observing the comparison between the cohorts with the 'original values' of ratings (first row), it can be seen that the rating patterns of anonymous and identifiable users looked differently. First, as previously outlined, with different predominant colours – green on Group 1 and purple on Group 2. Moreover, the 'first conversion' (second row) shows the distributions of 'likes' (in blue) and 'dislikes' (in red). This conversion can be used to infer, visually, the conformity that the users in each of the conditions experienced. For instance, when looking at the proportions of 'likes', it can be observed that, by being identifiable, people experienced greater conformity towards the quality of the posted questions, which is aggravated by using the dichotomous scale. Finally, the second conversion (third row) is presented to show that, even if 66% of the likert scale is converted into a 'like', Group 1 would still differentiate from the other two cohorts, thus demonstrating that the rating patters did change depending on the quasi-experimental conditions. Table 7.6 can also be seen as a summary of the three cohorts studied in the four chapters of results of this thesis.

Table 7.6 - Rating distributions per group, with original and converted ratings

	Group 1 (*N*=2,064)	**Group 2** (N=2,439)	**Group 3** (N=1,987)
Original values ■ 0 ■ 1 ■ 2 ■ 3 ■ 4 ■ 5			
1st Conversion Converting G1&2 to: [0, 1, 2] = Dislike [3, 4, 5] = Like ■ D ■ L			
2nd Conversion Converting G1&2 to: [0, 1] = Dislike [2, 3, 4, 5] = Like ■ D ■ L			

7.3.4 Understanding how different designs affect user's attitude towards the website

The intention of this last section of results is to determine how the attitudes towards being identifiable and towards the used rating scales: 1) correlated, and 2) affected the overall perception towards the website.

In Chapter 4, the attitudes were studied with a focus on the impact that being identifiable had on the website's perception. Similarly, in Chapter 5, the focus shifted to the effect of the attitude towards the rating scale on the website's perception. Conversely, this section analyses the effect that both attitudes – towards being identifiable and to the rating scales – had on the website's overall perception, using the data of the questionnaires from both Groups 2 and 3. This is done with the intention of understanding how all attitudes correlate, and also to obtain the percentage of a website's overall perception that can be explained by how helpful the users perceived its login requirements and rating scale. In order to do this, this section first presents a correlation among all the tested attitudes, where questionnaires took place. Thereafter, a categorical regression is offered.

Correlation among variables

As can be observed in the correlation tables below (Table 7.7), in both groups the variable that had the strongest correlation with the overall attitude towards PeerWise is how users perceived the adopted rating scale, which in both cases is significant at p<.001. The identifiability of users was not significant for Group 2 but is significant – and weakly correlated – for Group 3. In both cases, there is also a weak but significant correlation between the attitudes towards the rating scale and that towards identifiability.

Table 7.7 – Correlation between attitudes, Groups 2 and 3

Spearman's Rho Correlation Coefficient - G2	PeerWise's attitude	Rating scale's attitude	Identifiable usernames' attitude
PeerWise's attitude	1.000	.374**	.128
Rating scale's attitude	.374**	1.000	.204**
Identifiable usernames' attitude	.128	.204**	1.000

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Group 2

Spearman's Rho Correlation Coefficient - G3	PeerWise's attitude	Rating scale's attitude	Identifiable usernames' attitude
PeerWise's attitude	1.000	.329**	.245**
Rating scale's attitude	.329**	1.000	.206**
Identifiable usernames' attitude	.245**	.206**	1.000

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
a. Group 3

Effect on the website's overall perception

Finally, an ordinal regression was performed, in which the predictors were the attitudes towards being identifiable and to the rating scale, and the outcome variable was the overall perception of PeerWise. Using Cox & Snell's pseudo-R^2 value, results showed that in Group 2 the perception that users had regarding the helpfulness of the rating scale and being identifiable accounted for 22.0% of the variation of their overall attitude towards the website. This percentage was 26.2% for Group 3. The full results can be found in Appendix 7.2.

This last section offers a number of findings. First, the ordinal regressions indicate that about a quarter of the overall perception of PeerWise was due to the combination of the attitudes towards being identifiable and the adopted rating scale. Hence, this finding could be of use for SNS and CC designers (i.e. practitioners). Moreover, the correlations indicate that, from the two predictors, the rating scale had the most substantial impact on the website's perception, suggesting that a website's choice of a rating scale should be carefully planned. Further, and specifically for this research, it was interesting that, even in Group 2 where the emphasis was put on removing anonymity and making users identifiable, the attitude towards being identifiable was not even significant, whereas that towards the rating scale was significant and had the strongest effect on PeerWise's perception.

7.4 SUMMARY OF FINDINGS

This chapter targeted the overall research question, which aimed to understand the effect that different website designs (i.e. frames) affect the transmission of UGC. The presented qualitative and quantitative data showed evidence strongly suggesting that different frames do, indeed, produce affect the transmission of UGC by affecting: the authoring content (i.e. variation), which information to access (i.e. 'pre-selection), and how the UGC of others is rated (i.e. selection). This, in turn, affects the retained knowledge in the network; which again affects the variation and selection mechanisms.

The chapter encompassed four objectives. First, OB-2 was to understand knowledge transmission through the VSR mechanisms, making an emphasis on selection. In order to do this, the data from questionnaires and focus groups were used to investigate different ways in which participants relied on the network to make choices, and how these affected the transmission of UGC on each cohort. The qualitative data revealed that participants relied on the network at least in the following situations: for authoring, choosing, and rating questions. Regarding authoring questions (i.e. variation), students were given very little guidance on how to author the multiple-choice questions, nor were there written guidelines of what a 'good question' was. Hence, the comments from participants show that, by imitating others and through the feedback they received with the individual and overall ratings, each cohort arrived at a (non-spoken) agreement of what an 'acceptable' question looked like. In other words, this meant that the variation of UGC had been affected by the selection and retention mechanisms. Moreover, it was also discovered that pupils relied on the network for choosing questions, which could be seen as prerequisite for selection. Participants revealed to have targeted questions based on the knowledge that was available on the network (i.e. retention), such as the average ratings, perceived difficulty, or number of answers and comments, among others. Finally, regarding ratings (i.e. selection), users seemed to have relied on the summary of previous ratings (i.e. retention) to choose their own. Therefore, the comments of participants were not only useful to detect ways in which users had relied on the network, but also to unravel the different interactions between the VSR mechanisms.

Furthermore, the second tackled objective (OB-3&5) was to investigate how the different combinations of user profile and rating scales affected the transmission of UGC, by comparing the rating sequences from the three cohorts. Through the use of sequence lengths, transitions, entropy, and complexity, it was discovered that, even when the content is the same, website designs determine how users select information. Specifically, even when presented with the exact same information, each cohort arrived to a different mean rating, which arguably represented the groups' 'agreed' quality of the question. Moreover, it seemed that each cohort developed a unique 'rating pattern'. For instance, there appeared to be non-spoken rules (i.e. tacit knowledge) such as: 'use any rating in the scale and try to be critical/competitive' in Group 1; 'avoid extreme ratings and be nice/collaborative to others', in Group 2; and 'avoid the dislike rating at all costs – if you do not like the question just avoid rating it', in Group 3.

The third objective to be addressed (OB-7) involved deepening the study of conformity by comparing the 'levels of conformity' of each of the three quasi-experimental conditions. By making use of SA and transition rates, it was found that each cohort experienced different levels of conformity. Specifically, conformity towards positive values was of 60.8% for Group 1, 74.9% for Group 2, and 94.5% for Group 3. Hence, these percentages might give an estimate of how the different designs of SNSs and CCs can affect the choices of people. Namely, sites where users are allowed to sign-in anonymously and offer a broad rating scale (i.e. CCs), are expected to experience less conformity to the network. Conversely, sites where users are required to use their 'real' identities and use shorter rating scales (i.e. SNSs), are expected to experience considerably more conformity among members. Consequently, the results of this study agree with those from Lorenz et al. (2011), which determined that social influence (negatively) affects the 'wisdom of the crowd'. However, Lorenz et al. (2011) only determined that knowing the estimates of other members within the crowd affected individual's estimates; whereas the findings in this thesis found, in addition, that social influence also depends on *how* the estimates of others are presented. Specifically, the 'wisdom of the crowd' will be further undermined by having the estimates of others linked to their identities, and by a narrow scale used to give the estimates.

Lastly, the fourth tackled objective (OB-4&6) was to understand how the attitudes towards being identifiable and to the rating scale correlated, and the extent to which they could explain the overall perception of the website. Regarding this last point, it was discovered that about a quarter of the variation of a website's perception could be attributed to the attitude that users had towards the rating scale and being identifiable. Moreover, of these two, the one with the highest correlation were rating scales. Therefore, these findings suggest that the user-experience of a website is partly determined by the rating scale that the platform uses. Thus, it could be expected that online users will be very sensitive to changes in the rating scale because, as one participant put it, 'rating scales can cause problems of interpretation'. Moreover, knowing that users care about the design of the scales can have implications for practitioners, such as perhaps involving users before designing/changing a website's rating scale. Finally, given the effect that rating scales have on the choices of users and their web-experience, it could be argued that practitioners need to regulate them. The following two chapters – Discussion and Conclusion – will discuss all of these issues and their consequences thoroughly.

CHAPTER 8: DISCUSSION

This thesis has focused on providing a better understanding of how people within online networks choose UGC, through the analysis of two elements of website designs: user profiles and rating scales. Drawing on the findings of the previous four chapters of results and the conceptualisations put forward in Chapter 2, this section explores the impact of different website designs on decision making, through the relationship between processes of self-presentation and self-disclosure and group-level cognitive biases. Hence, this chapter discusses the key findings in the context of the reviewed theories, emphasising the '*what*' (i.e. the factors that should be considered in the explanation of the choice-making process), the '*how*' (i.e. how are the factors related), and the '*why*' (i.e. the psychological or social dynamics that explain the appointed factors) (Whetten 1989). Having thoroughly reviewed the choice-making process, the implications of choices (i.e. the selection mechanism) on the broader evolution of knowledge are then discussed.

Firstly, this chapter starts with a presentation of a processual model of choice making. After that, this process is unpacked by firstly exploring how websites seem to frame UGC through the set-up of user profiles and rating scales, and how these impact self-presentation and self-disclosure, and consequently the detection of groups and relationships. Secondly, the three group-biases are discussed, with an emphasis on investigating which had a more significant effect on users' judgement. Thirdly, conformity is addressed, with comparisons made between outgroups and ingroups, and ending with an examination of how this was affected by the different strength of ties. Finally, following this discussion which centres on the mechanism of selection, the chapter explores the implications of the research on wider evolving systems of knowledge within social media.

8.1 EXPLAINING THE SELECTION MECHANISM THROUGH CHOICE-MAKING

As previously mentioned, a salient contribution of this thesis is to explain selection providing a rationale based on choice-making. Moreover, as outlined in Chapter 2, this research proposed to study choices through the 'building blocks' of identity, groups and relationships (Kietzmann et al. 2012).

8.1.1 Discussion of the thesis' enhanced conceptual model

Drawing on literatures in identity, social psychology and interpersonal relationships (Kietzmann et al., 2012), this research contributes to the understanding of how choices are made within social networks. In Figure 8.1, shown below, the conceptualisation of the choice making process is represented. Choices, as presented by the ratings given by individuals, are based on a judgement they make following the interpretation of related information (see Figure 8.1). When making these judgements, individuals rely on heuristics to simplify the process of interpretation. This reliance becomes even more significant in complex environments, or when the individual is overwhelmed by inputs (as in the case of many online websites). However, it has been shown in prior research that the use of such heuristics can lead to systematic errors, or biases (Kahneman et al. 1982; Kahneman 2003). This study has focused on three core group-biases: *content* (i.e. specific qualities of a belief), *prestige* (i.e. successful individuals), and *conformity* (i.e. imitating the majority of the group) (Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011). These biases are in turn influenced by a number of factors, including the individual's identity (Goffman 1959; 1963), their membership of key groups (Tajfel & Turner 1979; Turner & Reynolds 2012), and their development over time of interpersonal relationships with other actors (Granovetter 1973; 1983; Cross, Parker, et al. 2001).

It is here that website design can affect biases through their influence on social processes, and the development of identity and interpersonal relationships. In this respect, social media websites are a good exemplar in which to study this process. They constrain the means in which individuals interact with others, regarding how individuals can present themselves, how much they can disclose, the type of content they can share, the number of characters that are allowed per post, and so forth (Kaplan & Haenlein 2010).

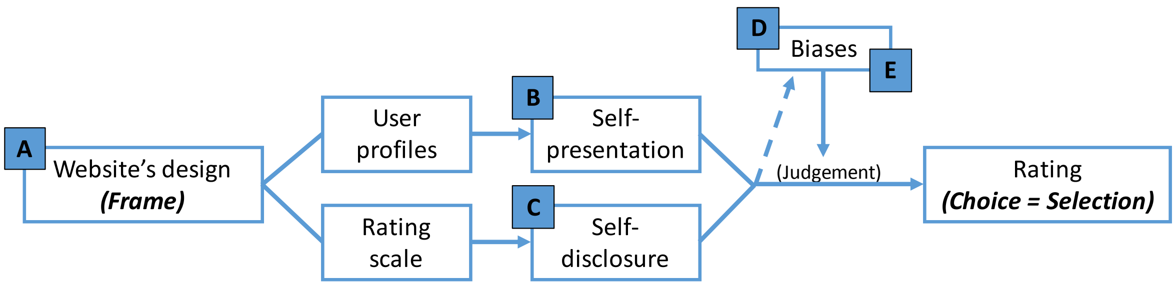


Figure 8.1 - An enhanced conceptual model of choice making within social media

The model proposed in Figure 8.1 reflects these influences in so far as they *frame* the way in which information is shared between individuals. These elements place boundaries on the ways in which individuals visually share information, thereby distorting the interpretation of information *vis-à-vis* the interactors. As a result, the beliefs (i.e. the replicators) derived from this UGC are correspondingly biased and distorted. As noted above, the study has focused on two aspects of a website's design: user profiles, which affect the levels of self-presentation of users; and rating scales, which modify the manner and degree in which people can express their likes and dislikes towards the posted content. Further, both self-presentation and self-disclosure shape the development of an individual's identity as those individuals interact with others through the development of interpersonal relationships (Goffman 1959; Goffman 1963; Tajfel & Turner 1979; Turner 1984). Hence, by influencing self-presentation and self-disclosure through user profiles and rating scales respectively, websites can *frame* the process through which choices are made by affecting the way in which individuals access and evaluate information (see Figure 8.1). By conceptualising this process through the dualism of interactor-replicator (i.e. UGC-beliefs), important interpretive processes are captured, highlighting the role played by frames on the wider evolution of information within the network.

The role played by these frames reflects the notion in prospect theory that the rationality of people is bounded by external and internal constraints, such as the way in which problems are formulated and individual's cognitive limitations (Simon 1955; Simon 1979). Regarding the construction of a decision problem, prospect theory established that the adoption of a frame has an impact on people's judgement, and therefore on their choices (Kahneman & Tversky 1979; Tversky & Kahneman 1981; Kahneman 2003). Therefore, frames have a significant influence on the evolution of information and knowledge between individuals and networks. By studying different website designs, this research allowed the exploration of the impact of different frames on this process through the different ways in which they enable or constrain the transmission of interactors (i.e. UGC). Manipulation of interpretative processes through different frames can thus shape the evolution of knowledge in social media (Figure 8.1).

8.1.2 Responding to the research questions through the enhanced conceptual model

The following sub-sections discuss the findings in relation to the research questions posed above. The relevance of each of these questions to the conceptual model is labelled in Figure 8.1. First, the concept of frame (A) is discussed, followed by the findings concerning the impact of user profiles on self-presentation (B), and the effect of rating scales on self-disclosure (C). Further, the results regarding biases (D) are commented on by highlighting the effect of the different levels of self-presentation and self-disclosure. Finally, the discoveries about conformity (E) are outlined, stressing the observed differences between ratings given to the whole network relative to those from users' personal groups.

A) The website's frame and its impact on the choice-making process of users

This part of the conceptual model refers to the overall question of this thesis: *How is the transmission of UGC affected by the different designs (i.e. frames) adopted by social media websites?* (see Figure 8.1, A)

As noted above, website designs influence the processes of self-presentation and self-disclosure which in turn affects the way in which individuals make choices. As outlined in Chapter 2, when rating content, individuals go through a cognitive process analogous to that of responding to a survey (Riedl et al. 2010). The process involves four steps: *comprehension* (attend to content and logically represent the question), *retrieval* (recall relevant information), *judgement* (integrate retrieved material and draw inferences based on accessibility), and *response* (map judgement into given options) (Tourangeau et al. 2000). However, it has been argued that individuals' rational choices are limited by internal and external constraints such as the way in which problems are formulated, their cognitive limitations, and the time available in which they should make the choice (Simon 1955; Simon 1979). In these terms, this research focused on the way information was formulated (i.e. its frame) and people's cognitive limitations (i.e. biases).

Tversky & Kahneman (1974; 1981) showed that the process of making choices partly depends on the formulation of the decision-problem (i.e. on how a problem was framed). Moreover, they observed that, when making a decision, individuals relied on heuristic principles (i.e. 'rules of thumb') which reduced the complexity of the task but would sometimes lead to 'severe and systematic errors' (i.e. biases). However, the shape and form of these biases depend on a number of social factors including; the identity of the individual, the group to which they belong and their relationships with the sender of information. The findings above identify clear links between the frames adopted in different website designs, and these social processes.

Much of the research done to date on social media is carried out within and not between sites (e.g. Wu et al. 2011; Wilkinson & Thelwall 2012; Liu & Park 2015; Park & Nicolau 2015; Jacobsen 2015). As a result, it is unclear the influence that different designs have on the choices of users, and ultimately on the evolution of content within these sites. Those few publications which do compare different platforms have instead focused on other issues such as different levels of gratification people obtain from them (Quan-Haase & Young 2010) or the personalities of their users (Hughes et al. 2012). By comparing platforms with the same degree of social presence and media richness (Kaplan & Haenlein 2010), different levels of self-presentation and self-disclosure could be examined. That is, it was decided to compare a CC and a SNS through the modification of a single online site so the users and the content on both were as similar as possible, and most of the changes could then be attributed to adaptations of the platform's design (i.e. its frame).

It was seen that the website's design (i.e. the adopted frame) influenced the transmission of UGC in two ways. First, the set-up of a site constrains the type and amount of content that users can post, and the manner in which they can respond to others. Consequently, this has an impact on how users express their beliefs, and on how other users interpret them. In evolutionary terms, this influences the relationship between the replicator (i.e. the belief), and its expression through the interactor (i.e. UGC). Therefore, poor website design can introduce significant interpretation issues, resulting in divergent choice-making processes. Second, websites also frame the content shared in the network by showing or concealing users' identities and therefore making them aware (or not) of who authored a particular post, and the relationship with the receiver of the UGC. This, in turn, influences the heuristics an individual will use when making judgements through content, prestige and conformity-based biases (see Figure 8.1). For example, people seem to perceive fewer differences with those among their ingroup, as they believe the errors in interpretation are less if they know the person sharing the information. As a result, they are more likely to conform to the views expressed within that group. Arguably, this makes people prone to perceiving as more valuable the content from those known to them, which also has an impact on their choices.

Furthermore, at the time when Tversky & Kahneman (1981) conducted their experiments on judgement, which consisted of presenting the same decision problem with different frames, they detected that people seemed to commit at least one of the following irrationalities: 1) They might choose a different option if the same question was framed in a different way. 2) They were normally unaware of alternative frames and the way these affected the attractiveness of options. 3) They would wish their preferences to be independent of the frame. 4) When they realised their choices were being inconsistent, they were often uncertain about how to resolve it. Hence, in line with Tversky & Kahneman's (1981) first outlined irrationality, it was found that participants made different choices depending on how the information available to them was framed. That is, the analysis of online interactions showed that people would give different ratings – even when the posted content was identical – depending on whether the author of the question was anonymous or identifiable, and whether the rating scale was likert or dichotomous. Following an examination of questionnaires and focus groups, it was further seen that the majority of users did not recognise that the adopted frame affected their own ratings, supporting Tversky & Kahneman's (1981) second proposition. Consequently, the findings from this thesis' quasi-experiments provided evidence to support the view that at least two of the proposed irrationalities took place, therefore proving that online environments frame UGC through modifying the levels of self-presentation and self-disclosure available to users.

To conclude, the overall research question can be answered as follows: the frames adopted by websites influence the choices made by individuals through their impact on choice making heuristics. Hence, users presented with the same information on a CC and a SNS site would most likely rate it differently. This study is the first attempt to investigate if different types of social media distinctly frame UGC, thus influencing the way in which individuals make choices. It should not be forgotten that "the adoption of a decision frame is an ethically significant act" (Tversky & Kahneman 1981, p.458). As a result, demonstrating that the set-up of a website affects the choices of people and bounds their rationality, has a number of theoretical, practical, policy, and ethical implications which are discussed thoroughly in Chapter 9.

B) The effect of user profiles in self-presentation: anonymous vs identifiable

This part of the conceptual model relates to the first research question of this thesis: *How are the choices of users affected by different levels of self-presentation occurring from diverse user profiles?* (see Figure 8.1, B)

This question explores in more depth the role played specifically by self-presentation in the choices made by individuals. In offline environments, self-presentation has been defined as a component of identity by which an individual tries to make an impression on others (Goffman 1959; 1963). This impression is subject to elements that the individual cannot change, called 'personal front' (e.g. gender, age, racial characteristics), components that are under her/his control named 'setting' (e.g. clothing, personal adornments), and the person's 'performance' (i.e. the actions or behaviour at a given occasion which serve to influence others), (Goffman 1959). However, online environments combine the characteristics of disembodiment and anonymity, which result in a new means of identity construction (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008). The internet gives individuals the opportunity to alter their identity to the extent that would not be possible in face-to-face communication, by allowing them to change their age, gender, appearance and personality (Suler 2002). However, this opportunity to alter several aspects of the self or even remain anonymous does not occur in the same manner in all social media sites.

For instance, SNSs (e.g. Facebook) are classified as having high levels of self-presentation, whereas CCs (e.g. YouTube) are seen as having lower degrees (Kaplan & Haenlein 2010). As described in Chapter 2, the profiles of CC users require very little information and, in most cases, individuals have the option to remain anonymous (see Figure 2.2). Conversely, user profiles of SNSs require more personal details (e.g. Figure 2.1), and users are highly encouraged to sign in with their 'true' identities[34], which is why these sites are called *nonymous*, or non-anonymous (Zhao et al. 2008). For this reason, the present study used a single website in which participants in the first cohort were anonymous, while later cohorts were non-anonymous (i.e. identifiable). It was observed that, by altering this element of a website's design (i.e. the absence/presence of user profiles), the level of self-presentation was affected and this, in turn, modified the choices

[34] That is, with an identity which is already known and recognised by others.

of users as reflected through their ratings. Specifically, ratings were affected in two ways: 1) when anonymous they seemed to behave competitively, whereas when identifiable they behaved collaboratively; and 2) the so-called 'personal front' became significant when users were identifiable. Drawing on literature of social identity the mechanisms behind these findings can be further explored.

First, the most remarkable finding regarding self-presentation was that participants in each experimental condition adopted what could be described as a different strategy, whereby anonymous users were seen as acting 'competitively', whilst identifiable ones seemed to adopt a 'collaborative' orientation. Similar findings have been found by other scholars but, depending on the area of study, it has been described as effecting positive or negative outcomes. For instance, an area where anonymity is seen as having a positive outcome is education. Recent educational studies utilising peer-review with computer-mediated communication (i.e. e-peer review), have found that anonymous groups give more 'critical' feedback, leading to better performance (Lu & Bol 2007; Howard et al. 2010; Li 2017). Conversely, an area where anonymity is mostly seen as having a negative effect is cyberbullying. Researchers have found that bullying tends to happen at higher rates in sites where users are anonymous (Whittaker & Kowalski 2015; Brandtzæg et al. 2009). Nonetheless, it should be mentioned that there is a smaller group of researchers studying cyberbullying who claim that anonymity is more a fear than a reality; most cyberbullies are not anonymous and are rather part of the victims' real-life group (Mishna et al. 2009; Huang & Chou 2010).

Regardless of whether scholars see anonymity as having a positive or negative outcome, all agree that it changes behaviour. There are mainly two underlying processes which they have used to explain this issue. On the one hand, some base their findings on *deindividuation theory* which explains group behaviour through the assumption that there is only one concept of self, which is reduced and subject to the group norms when an individual is within a crowd (Le Bon 1896; Zimbardo 1969). Deindividuation theory is mostly used as a justification for anti-normative behaviour, so the online research that makes use of it argues that users lose their 'self' to the network, and this situation is aggravated with anonymity and the use of computer-mediated communication (e.g. Lu & Bol 2007; Brandtzæg et al. 2009). On the other hand, others have relied on the *social identity model of deindividuation effects* (SIDE), which argues that 'the self' is not lost when in a crowd because the self-concept is instead something that adapts to diverse contexts, groups, and situations (Lea & Spears 1991; Reicher et al. 1995; Postmes et al. 2001). Researchers justifying interactions through SIDE do recognise that in some situations crowds can have negative behaviours, but also argue that anonymity within groups can generate altruistic acts among group members, although this would depend on the norms and beliefs of the in-group (e.g. Howard et al. 2010; Whittaker & Kowalski 2015; Mishna et al. 2009; Douglas & McGarty 2001).

Considering these two approaches, the findings of this present research reflects the SIDE perspective for three reasons: firstly, as has been mentioned throughout the research, it is assumed that individuals will vary their 'self' depending on the audience and context (Goffman 1959; Goffman 1963). Secondly, SIDE is based on the assumption from social identity theory that the self-concept varies in a continuum between the individual (i.e. interpersonal) and as members of a social group (i.e. intergroup) (Lea & Spears 1991). As outlined in Chapter 2, this is also one of the suppositions of this research. Given that individuals easily cluster into groups, they thrive for a positive association with their in-groups, as they struggle to maintain a positive self-concept. However, their behaviour towards members of the in-group and out-group would be determined by their perception of the social setting, as either interpersonal or intergroup (Tajfel 1974; Tajfel 1978; Tajfel & Turner 1979). Thirdly, SIDE is also grounded on the self-categorisation principle (Reicher et al. 1995), which assumes that individuals define their identities through minimising their differences with other members of the ingroup (Turner 1984; Turner & Reynolds 2012). Likewise, 'competitive' and 'collaborative' strategies – observed when users were anonymous and identifiable, respectively – can be explained through the three outlined assumptions, with one alternative perspective on the second point.

The alternative perspective concerns the interpersonal-intergroup continuum. As outlined, Tajfel (1978) and Tajfel & Turner (1979) assume that there is a continuum between interpersonal (i.e. personal) and intergroup (i.e. social) behaviours. However, Stephenson (1981) argued that there were situations where both interpersonal and intergroup components would affect situations. In his view, there would be situations where "both interpersonal and intergroup goals may be strong; these may often conflict, and yet a choice between the two may still have to be made […]. It cannot be assumed that [individuals][35] perceive settings in interpersonal or intergroup terms or seek interpersonal or intergroup goals in an inverse ratio, simply because in some instances they may be forced to act in this manner" (Stephenson 1981, p.193). Therefore, rather than a continuum, he proposed mapping interpersonal and intergroup relations as independent dimensions. Figure 8.2 shown below presents a graphic representation of how the continuum was proposed to be transformed into independent dimensions (Stephenson 1981, fig.6.4 & 6.5, p. 190-197).

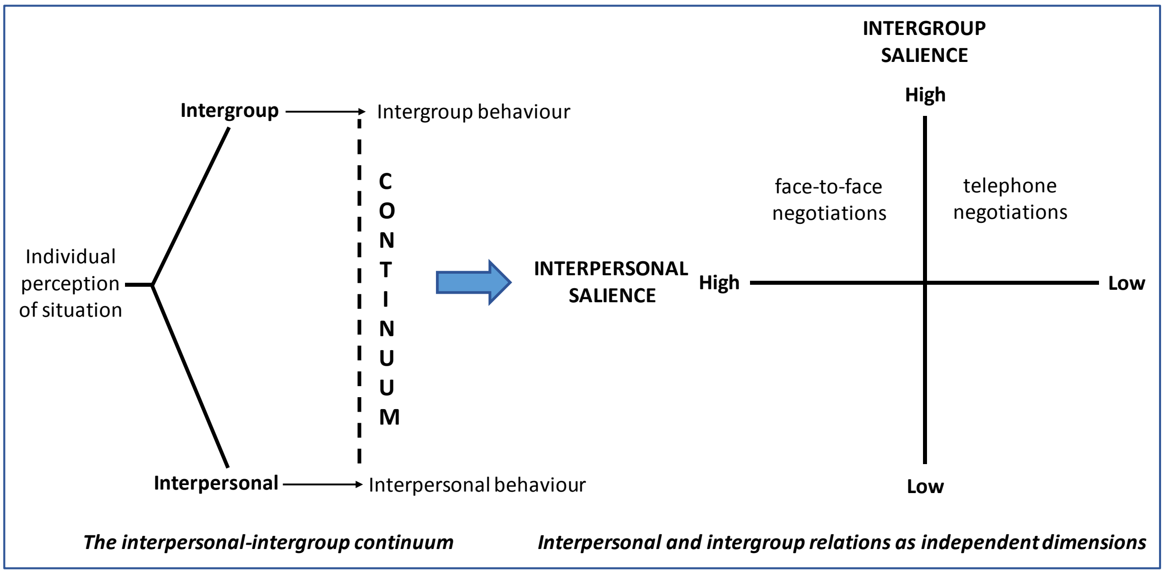


Figure 8.2 - Interpersonal-Intergroup continuum & dimensions (Stephenson 1981, pp.190-197)

Further, from the figure shown above, it is seen that different means of communication (face-to-face and telephone) produced different outcomes regarding interpersonal salience. Similarly, given that social media was defined as a 'communication system' (Peters et al. 2013), it could be argued that different website designs can produce different degrees of interpersonal and intergroup salience. Figure 8.3, below, maps the different interpersonal and intergroup degrees corresponding to the three distinct scenarios studied in this research.

[35] The original quote reads 'negotiators' instead of 'individuals'

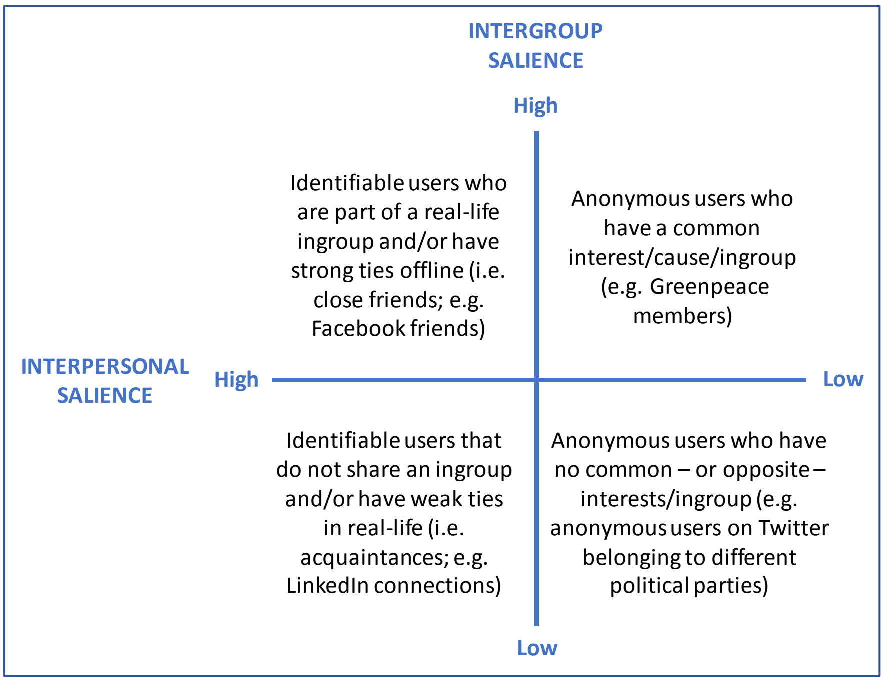


Figure 8.3 - Interpersonal and intergroup dimensions of the different designs of user profiles

In sum, considering these dimensions of interpersonal and intergroup salience (see Figure 8.3), the findings can be interpreted in the following manner. When website design constrains an individual's ability for self-presentation, resulting in anonymity, actors are not assimilated into group norms (i.e. self-categorisation is low), leading to them behaving in a more individualistic manner, with high interpersonal salience. Nevertheless, because they cannot determine who the other users are (i.e. the interpersonal dimension is also low), they base their ratings on what is available: content, which is why peer-reviews are more objective when users are anonymous. Conversely, when identifiable, the intergroup characteristics (i.e. ethnicity, gender, group membership) become relevant, which can explain in the current study: 1) why users accessed more UGC from those similar to them; and 2) why ratings overall were higher and more similar, as users went through a process of self-categorization with their in-groups. Further, because all users were able to recognise some people they knew outside the website, the interpersonal dimension became relevant in these situations, which resulted in ratings being almost proportional to relationships.

To conclude, the answer to the first research question is: the choices of users are significantly affected by the different levels of self-presentation generated by the use of the characteristic profiles of CCs and SNSs. Hence, ratings differ when users are anonymous or identifiable. This difference happens because, by being identifiable, both the intergroup and interpersonal dimensions of situations become relevant; and, when this happens, users tend to self-categorise with their ingroups and therefore acquire more beliefs from them.

C) The effect of rating scales in self-disclosure: likert vs dichotomous

This part of the conceptual model concerns the thesis' second research question: *How are the choices of users affected by different levels of self-disclosure happening due to the use of distinct rating scales?* (see Figure 8.1, C)

The impact of self-disclosure, defined as the revelation of personal information such as opinions, likes and dislikes (Kaplan & Haenlein 2010), was explored through the use of two types of rating scales: likert and dichotomous. As was mentioned in Chapter 2, there seems to be contradicting instructions regarding the use of broad or narrow scales. On the one hand, researchers advocate for the use of wider (i.e. more granular) rating scales (Riedl et al. 2010; Riedl et al. 2013). On the other hand, practitioners are adopting and recommending the use of narrower scales (Ciancutti 2011; YouTube 2009). The results of this research seem to have a stronger agreement with those of scholars, for reasons that are discussed below.

Self-disclosure impacts directly on the interpretive processes which underpin choice making. As discussed above, the website design shapes the broader evolution of UGC by influencing the interpretive processes through which individuals make choices to select UGC, and their corresponding beliefs. The different designs of CCs and SNSs constrain both the type and amount of content that users can post (as senders of information) and the manner in which they can respond to others (as receivers of information and givers of feedback). Subsequently, this has an impact both on the way users express their beliefs (replicator to interactor), and on how other users interpret them (interactor to replicator). Through their influence on interpretation, degrees of self-disclosure thus influence the choice making process.

In this study, rating scales have been used to indicate the choices of users; that is, to map their judgement towards a particular piece of UGC. Yet, this evaluation conveys a belief itself, such as: 'The content is good', 'I do not like this video', 'This post is helpful', and so forth. So, from a communication viewpoint, ratings can be seen as a way of responding or giving feedback to the person sharing the UGC. Then, the 'sender' (i.e. the individual who authors the information) also needs to interpret the feedback given by the 'receiver' (i.e. the individual who rates the content). Moreover, given that a characteristic of UGC is that it must be publicly available online (OECD 2007), other users in the network must also interpret the observed communication between the sender and the receiver, given that they might want to know the choices of others before making their own. Therefore, the design of rating scales can thus restrict the way in which users can respond to others, as it determines the number of options that the receiver has to express feedback visually (i.e. the interactor). This has three consequences. First, the restrictions on the visual way in which feedback can be given has an impact on how the receiver can express her belief (i.e. replicator) regarding the evaluation of the content s/he observed. Second, it affects the way in which the sender interprets the feedback she gets from the receiver. Third, it also influences the way in which other users observing this communication make sense of it.

The results from the two experimental conditions – likert and dichotomous rating scales – can be used to illustrate these three points further. First, regarding the effect that scales have on the person giving feedback, Chapter 5 showed that three-quarters of the people using the likert scale found it helpful or very helpful, whereas this satisfaction dropped by 20 percentage points for those using the dichotomous scale (see Figure 5.10 and Figure 5.11). This evidence, together with the insights provided by participants, showed that they felt restricted by having only two options to express their opinions (i.e. their beliefs), and accordingly determined that they would instead not rate. Hence, adopting a specific rating scale can affect the way in which people self-disclose, which has an impact on their satisfaction and can make them more (or less) prone to use the rating scale.

Second, regarding the impact of the scale on the sender's interpretation, there were two discoveries. 1) Although the users of both scales complained about issues of interpretation (e.g. not understanding why someone had given them a specific rating), it seemed to be aggravated by the use of the dichotomous scale. Moreover, 2) several users mentioned that there appeared to be an 'emotional' meaning attached to the 'thumbs up/down'; specifically, 'dislike' was seen as 'mean' and was therefore avoided (see Table 5.3). These two points thus illustrate how certain rating scales designs can create greater (or lesser) interpretation issues, and can even involve more sentiment than others. Third, an example of the influence that rating scales can have on other users in the network can be found in Chapter 6, where users from both cohorts mentioned they would look at the average rating of questions before tackling them. Once again, this was intensified by the use of the dichotomous scale, as every 'dislike' would have a greater effect on the average-value of questions because it had assigned a value of zero. Consequently, questions that had even one 'dislike' were less likely to be accessed by users in the network.

Previous research conducted on rating scales found that people using a likert scale experienced greater satisfaction when rating compared with those utilising a dichotomous scale (Riedl et al. 2010). The explanation for this was that the latter forced people into making a binary decision, which produced higher levels of stress (Riedl et al. 2010). However, the cited experiments were carried out under conditions of anonymity and isolation. Conversely, in this thesis, the different scales were studied in the context of identifiable users. Consequently, self-disclosure is aligned with self-presentation. Therefore, a plausible explanation for users feeling more restricted by the dichotomous than the likert scale is that the stress of making a decision would be increased because they were in a social context. Therefore, given that one option of the scale had a positive connotation while the other had a negative one, they would avoid rating in a manner which could lead to online or offline retaliation. Finally, it was discovered that senders (i.e. authors) interpreted the positive ratings as fairer, reflecting the need for self-esteem and reward (Tajfel & Turner 1979).

To conclude, answering the second research question: this research has found quantitative and qualitative evidence that confirms that the choices of users are affected by the different rating scales used online. This impact is caused by the effect that rating scales have on the interactor-replicator, because of transmission errors in communication, by different levels of stress generated while making a choice, and as a consequence of psychological and physiological effects of rating scales on groups and individuals.

D) The effect of different degrees of self-disclosure and self-presentation in biases

This part of the conceptual model (see Figure 8.1, D) concerns the study of the three outlined group-biases: content, prestige, and conformity (Henrich 2004; Richerson &

Boyd 2005; Mesoudi 2011). This section is not linked to a research question because an emerging finding of this research was that the hierarchy and strength of these biases were affected by the different levels of self-presentation and self-disclosure.

The finding noted above points to important differential effects between different levels of self-presentation and self-disclosure and the hierarchy and strength of these biases. To be precise, depending on the different scenarios of user profiles and rating scales, both the type of group-bias that was more prevalent and its intensity were affected. To explain this finding, each scenario of the quasi-experiment will be briefly discussed in relation to group biases.

In the first quasi-experimental condition, users were anonymous and utilised a 0-5 likert scale. This setting resembled the design of most CCs, as explained in Chapter 2. Under these circumstances, the predominant biases were seen to be prestige-, followed by content-based, as outlined in Chapter 4. Moreover, as users were anonymous, they knew very little about their 'real-life prestige' (except recognising the 10 questions posted by the module 'admin'). Hence, they needed to rely more on the 'online-gained prestige' earned online through the built-in gamification elements (e.g. badges, reputation scores, and previous ratings given by users). These results are similar to studies regarding online reviews, where it has been found that elements such as the reviewer's online reputation and the number of fans or followers had a significant positive impact on other users' perception of usefulness towards their posted content. In addition, these prestige-related variables had a more significant impact than those which could be catalogued as content-based (e.g. readability index, word count) (Cheng & Ho 2015; Liu & Park 2015).

Other studies on real-life prestige (i.e. not in an online context) have found that when information on individuals was not available, status was inferred: 1) through observing the behaviour within a group; or 2) from the judgement of others, which could be coded through 'markers' such as university degrees and awards (Henrich & Gil-White 2001). Therefore, given that online environments allow for the identity of individuals to be concealed to a greater extent than in offline settings (Douglas & McGarty 2001; Bargh et al. 2002; Suler 2002; Suler 2004; Zhao et al. 2008), it is understandable that people would rely more on the observed ratings of others and on the available online symbols that denote prestige. Reflecting on the findings of this first study, it can be surmised that individuals tend to follow the lead of others who are assumed to have more 'prestige'. Biases allow individuals to economise on valuable cognitive resources when making decisions (Kahneman et al. 1982), and in the absence of clear group allegiances – as a result of anonymity – individuals herd around perceived shepherds.

In the second quasi-experimental condition, users were all identifiable, while continuing to use the likert rating scale. This design is similar to those CCs that require users to login with their SNS profiles. With this second arrangement, the strongest group-bias was conformity, followed by content-, and finally prestige-based, as outlined in Chapters 5 and 6. Thus, it can be seen that, by modifying the design of user profiles and therefore increasing levels of self-presentation, the hierarchy of the biases changes compared with the first scenario. As mentioned in Chapter 2, conformity has not yet been explored in online settings through the study of real-life relationships. However, a handful of studies on eWOM have found a positive relationship between a review's perceived usefulness and increased levels of self-presentation by users (i.e. 'identity disclosure'), which have been studied through elements such as the author's use of her real name, profile picture, and address (Liu & Park 2015; Park & Nicolau 2015). Unfortunately, when extracting online data, the information regarding which users gave useful or helpful votes is lost. It would be interesting to see if these votes were given because authors of the reviews displayed higher levels of self-disclosure, or because this allowed similar individuals to identify with the author of the review.

The fact that conformity-based biases became the prevalent heuristics in the second study can be explained in terms of the social and interpretive processes discussed above. Evolutionary scholars have claimed that individuals evolved through imitating others within the population, as this was quicker and less risky than figuring out things by themselves (Henrich & Boyd 1998). For instance, eating whichever food was consumed by the members of a group was less risky than trying out food by themselves, which may turn out to be poisonous (Mesoudi 2011). Therefore, conformist transmission is responsible for group members holding analogous beliefs and behaving similarly, consequently creating and maintaining group boundaries through time (Henrich & Boyd 1998; Richerson & Boyd 2005). Similarly, as has been discussed throughout the thesis, when characteristics such as gender, ethnicity, and group memberships are visible, they cause people to align their identities and favour their ingroups (Tajfel 1974; Tajfel &

Turner 1979). Hence, allowing individuals to identify each other online makes them more prone to acquiring content and, presumably, beliefs from those who are part of their real-life in-groups because, from an evolutionary and social psychology perspective, this is how individuals have learnt to acquire knowledge from groups. By increasing degrees of self-presentation and self-disclosure, the second study thus allowed individuals to develop identities and intergroup tendencies (see above), resulting in conformity-based biases coming to the fore.

Lastly, in the third quasi-experimental condition, while users remained identifiable, the rating scale was changed to be dichotomous. This setting is actually how most SNSs are set-up. In this last scenario, discussed in Chapter 5, the hierarchy of the type of group-biases remained the same as in the previous set-up, but the strength of conformity increased, as seen through the sequence analysis in Chapter 7. Therefore, the change in the level of self-disclosure that occurred as a consequence of modifying the rating scale altered the intensity of the group-bias. As discussed above, the restricted degree of self-disclosure found from using the dichotomous scale both constrains the individual's ability to express and interpret beliefs. By limiting the interpretive processes in this way, individuals searched for cues from others and conformed to their behaviours. Moreover, the higher degree of stress associated with negative outcomes of this rating scale (Riedl et al. 2010), acted to further increase these tendencies towards conformity.

To sum up, by altering degrees of self-presentation and self-disclosure, both the hierarchy and intensity of group-based biases changed. Individuals are boundedly-rational creatures (Simon 1955; Simon 1979), who fall back on simple heuristics to make choices under conditions of uncertainty and complexity (Kahneman 2003). These results highlight the critical effect that different website designs have on networked individuals, as diverse setups would determine – to some extent – the prevalence and strength of the biases within the groups. It is worth remembering that biases were defined as 'severe and systematic errors' (Kahneman 2003). Such errors undoubtedly shape the broader evolution of knowledge and content within such networked environments. Thus, showing that the way in which online environments are designed can lead to more acute systematic mistakes is of relevance and has several implications that will be further discussed.

E) The effect of the presence of personal networks in online environments

The last part of the conceptual model concerns the third research question: *How are the choices of users affected by the online presence of their personal networks?* (see Figure 8.1, E)

One of the most relevant contributions of this thesis has been to propose, justify, and demonstrate the importance of differentiating between conforming to the whole network and users' personal networks (i.e. those also known to the user in real-life). As previously mentioned, this differentiation would not be necessary for studies that assume all group members know each other (e.g. Henrich & Boyd 1998; Henrich 2004; Richerson & Boyd 2005; Mesoudi 2011; Perreault et al. 2012). However, it has been argued that this should be an essential distinction in research concerning online environments, especially in those studies that focus on reviews or ratings (e.g. Riedl et al. 2010; Riedl et al. 2013; Liu & Park 2015; Cheng & Ho 2015; Park & Nicolau 2015; Mueller et al. 2018). It is certainly difficult to obtain the personal networks of online users but, given the high number of websites and apps that are encouraging people to sign-in with their SNS accounts, it has become a necessity to understand the effect that these are having on the transmission of information. Furthermore, it should be highlighted that the current research is the first attempting to differentiate and measure the impact that both the whole and personal networks have on people's choices, in an existing online environment and within a real social setting.

Regarding personal networks, Chapter 5 showed how – as soon as users were identifiable and regardless of the used rating scale – individuals favoured those known to them by giving them higher ratings (see Figure 5.6 and Figure 5.7). Further, Chapter 6 explored different cluster arrangements, based on binary values of friendship (i.e. absent/present) and discovered that regardless of the size of the cluster, users would still favour their ingroups (see Table 6.4). As outlined in the previous section, this tendency is due to the related impacts of both social and cognitive psychological factors. While the latter highlight the role played by cognitive heuristics in decision making; the former refers to the impact of social forces impacting upon the development of identity within a social group (Tajfel & Turner 1979). The tendency towards conformity biases in turn "creates and maintains group boundaries and cultural differences through time. Such boundaries may establish the initial conditions that lead to the development of group stereotypes, ethnic conflict, and racial strife" (Henrich & Boyd 1998, p.231).

Perhaps the dichotomous differentiation between in-group and out-group would have been enough less than a decade ago when social media used to "treat all users the same" (Gilbert & Karahalios 2009, p.211). However, as discussed in Chapter 2, since 2011 SNSs started to differentiate among different strengths of relationships (Vahl 2011; Loomer 2012); and this rolled-out to CCs in 2013 when the latter started suggesting that users log in with their SNS profiles. For this reason, it was decided that the study would benefit from making a further differentiation of individuals' personal networks. In this regard, scholars in the field of social networks found that a person's information environment consisted – in large part – of the relationships they had (Cross, Parker, et al. 2001). However, the transmission of information was affected by the strength of ties, which are determined by the time and intimacy of relationships (Granovetter 1973; 1983). Thus, this study differentiated between *absent* (i.e. no relationship), *weak* (i.e. acquaintances), *intermediate* (i.e. friends), and *strong* ties (i.e. close friends). Literature dictates that the importance of strong ties rely on the transmission of complex knowledge and collective beliefs (Dobson et al. 2013), whereas weak ties are essential for the diffusion of information (Granovetter 1973).

Regarding the results from the quasi-experiment concerning ties, it was found that 9 out of 10 ratings were performed among users that had absent ties (see Table 6.3). However, people rated the content of those among their personal networks consistently higher. Specifically and remarkably in all cases, the average ratings were proportional to the closeness between users, or the strength of their ties. Therefore, these findings suggest that online users do not necessarily try to access specifically the information from those in their personal networks. However, when they come across or purposefully find it, they do perceive that content as being of better quality, or at least they act as if it were of higher value. That is, in accordance with the theory of strength of ties (Granovetter 1973; 1983), people probably realise that only acquiring UGC from their closer connections would deprive them of information from distant parts of the social network. However, given that ratings among close friends are also higher further confirms that stronger ties hold collective beliefs (Dobson et al. 2013). In brief, individuals seek out groups of like-minded individuals within the broader network of connections. They increasingly draw on the interpreted beliefs of these groups as they make choices. Given the influence of website frames on antecedent social processes, the choice of designs thus has a compounding effect on the wider evolution of content within the network.

There could be two explanations as to why ratings were proportional to the strength of ties, both related to identity. Firstly, if it is assumed that ratings are an accurate representation of the evaluation of content, higher values would reflect a tendency towards conformity within local groups. The emergence of local groups and the differentiation of ratings among those within ingroups once again contradicts the assumption of the interpersonal-intergroup continuum, which assumes that all in-group members would be treated alike (Tajfel 1978; Tajfel & Turner 1979). Equally, the continuum assumes that all out-group members are treated in an undifferentiated way which, as could be seen from the geodesic distances (see Table 6.5), was also not the same. Therefore, these results further support Stephenson's (1981) proposal of independent intergroup-interpersonal dimension (see Figure 8.2). Consequently, the closer the relationship between a group, the more individuals will tend to self-categorise and conform, which translates into higher perceived UGC quality and higher associated ratings.

Secondly, it is possible that higher ratings were not reflective of the 'true feelings' of individuals. Drawing on the qualitative findings gathered from questionnaires and focus groups, it was seen that some of the ratings were given as a result of 'social pressure' as opposed to perceived quality. This conformance to the perceived group-level assessment of quality reflects *embarrassment* avoidance as actors seek to behave in a manner consistent with their 'official projection' (Goffman 1959). 'Online embarrassments' can even lead to users presenting different identities on diverse sites, or they can occur when they present a different identity online from offline (Kietzmann et al. 2012). By being identifiable, individuals thus seek a level of congruency between their online and offline identities. This could have been translated into acting 'nicer' through giving higher ratings to those with whom they had stronger ties. However, at the individual level, the truth probably lies somewhere between acting and truthfully categorising with their in-groups. As mentioned throughout the results, ratings seemed to be a combination of these two scenarios: sometimes they really represented the evaluation of the content, and in other circumstances, they were more of a conscious act to please their friends. In either situation, at the network level, the result was that people showed higher conformity to their closest groups.

Lastly, and in relation to the debate on social networks, it is worth reflecting on the notion of collective intelligence. The term ‘wisdom of the crowds’ has been used to describe how groups can be smarter than the single individuals comprised in them (Surowiecki 2004). Further, some researchers have referred to it as ‘collective intelligence’, using examples such as Google and Wikipedia as proof that networks can outsmart individuals (Malone et al. 2009; Malone et al. 2010). Specifically, a handful of scholars have even argued that UGC, in the form of reviews and votes, can be used to “automatically generate remarkably accurate verdicts” (Hill & Ready-Campbell 2011, p.73). Conversely, other scholars have been more sceptical about this issue, and have highlighted some of the limitations of crowds (e.g. Roman 2009; Bonabeau 2009). Moreover, one study concluded that social influence would actually decrease the intelligence of the crowd, claiming that groups were initially wise, but knowing the estimates of fellow participants within the experiment “narrowed the diversity of opinions to such an extent that it undermined the wisdom of the crowd effect” (Lorenz et al. 2011, p.9020). Two reasons might be put forward to support their claim: 1) social influence diminished the diversity in the network, without improving its accuracy; and 2) the convergence of opinion increased people’s confidence, despite the lack of improvements.

The findings of this study show similar results to that of Lorenz et al. (2011). Chapter 7 presented a comparison of the levels of conformity attained through the three quasi-experimental conditions, seen through the transition rates of sequences. As seen in Table 7.5, the conformity towards the group went from 61% when users were anonymous and utilised a likert rating scale, to 75% with identifiable users and a likert scale, and finally to 95% when users remained identifiable and the rating was changed to dichotomous. Further, the reasons behind these findings might be explained drawing on the discussions given by Lorenz et al. (2011). Conformity is known for having a significant ‘side effect’, which is reducing the amount of variation within groups (Richerson & Boyd 2005). Moreover, as outlined in Chapter 2, individuals within a group try to self-categorise by minimising the existing differences between them and other in-group members (Turner 1984; Turner & Reynolds 2012). Therefore, by increasing self-presentation and reducing self-disclosure – that is, by going from a CC to a SNS design – conformity and self-categorisation increased, translating into less variation and stronger common beliefs, which made people converge towards the same ratings. Therefore, the importance of website design should be highlighted once again, as this can affect levels of conformity within the network. The impact of these processes not only on the mechanism of selection, but also the mechanism of variation has significant implications for the broader evolution of content within the network (as discussed further below).

Finally, to summarise the answer to the third research question, users showed higher levels of conformity towards their personal networks. Specifically, ratings were proportional to the strength of their ties. At the individual level, this could be explained by a mixture of social identification with the group, and avoiding real-life embarrassments with friends. Nevertheless, higher conformity can translate into a fewer variation of beliefs and less smart choices by the crowd. Thus, social media sites that exploit people's personal networks are more likely to produce higher conformity within members of the network.

8.2 THE EVOLUTION OF KNOWLEDGE WITHIN SOCIAL MEDIA

Whilst the discussion above has focused on the conceptualisation of the selection process, and the impact that different website designs have on this, there are also implications for the wider evolution of knowledge within the network. To explore these processes further, the network has been conceptualised as an evolving system in which UGC evolves over time through the VSR mechanisms defined by Campbell (1960). As noted in the earlier chapters of this thesis, this research has taken a Generalised Darwinist position which generalises the VSR mechanisms to fields such as sociocultural evolution (Breslin 2010; Breslin 2011). This stance assumes that there is a pool of 'replicating entities' (Hodgson 2005), conceptualised through the dualism of the replicator and its developmental expression, the interactor (Dawkins 1976). The *variation* mechanism supposes that the replicators are sufficiently different from each other to be distinguished and later selected; *selection* implies that the units that evolve go through a competition in order to be chosen; and *retention* assumes that the entity that was selected (i.e. the replicator) can be identified to be kept by those who have selected it, or can also be 'inherited' intact or with modification (Mesoudi 2011).

This research has advanced the understanding of the process of knowledge evolution by theorising and investigating how the VSR mechanisms take place in online environments. In addition, it has extended the study of the selection mechanism, both in online and offline settings, by providing a rationale based on choice-making. It should be noted that already the literature review of this thesis contributed to theory by hypothesising how the VSR mechanisms would take place in social media and how knowledge would be translated into information and vice versa. However, thanks to the quasi-experiments conducted, it has been possible to reflect on what was initially theorised and make further improvements to what was predicted in Chapter 2 of this thesis.

8.2.1 Conceptualising the VSR process at the individual level

First, as discussed in Chapter 2, it was argued that content evolves in online systems through the variation of an individual's beliefs (i.e. the replicator), the subsequent selection of related UGC (i.e. the interactor), and finally the retention of interpreted beliefs of the other. Prior to the completion of the quasi-experiments, it was assumed that all beliefs would be coded into information and therefore always visible to users and the researcher. However, after conducting the quasi-experiments and analysing the data, it became evident that there was some knowledge that was not explicitly mentioned; that is, it was only tacit (Polanyi 1968; Nonaka 1994). This discovery was made not as a result of the information retained by the network, but because there was a noticeable change in the way that participants from each cohort interacted. For example, when using the dichotomous scale not giving a 'like' was a way of communicating that the content was not considered adequate. Therefore, in these situations, knowledge was transmitted even without the use of interactors and was hence not observable. To further explain how knowledge was sometimes transmitted without being coded into UGC, it is worth drawing a comparison with offline communication.

For instance, regarding face-to-face discussions, it is well documented that verbal messages comprise only a small percentage of interpersonal interactions (Forgas 1985). It was actually Darwin (1872) who began this research by comparing the expression of emotions in humans and animals. After almost a century, the study of *non-verbal communication* (NVC) surged as cross-disciplinary research among psychiatrists, linguists and anthropologists (Hecht & Ambady 1999). NVC comprises signals like facial expressions, posture, eye-movement, gesticulations, and tones of voice (Argyle et al. 1971). Moreover, it has been argued that these signals affect interpersonal communication, feedback, and self-presentation, and are therefore essential for social interaction (Argyle 1969). NVC has been so relevant in the social domain that some social psychologists argue that it should be called *non-verbal behaviour* (NVB), claiming that even facial expressions reflect previously adopted behaviours (Krauss et al. 1996). It is possible that Darwin would agree with the latter statement, given that he used words like *action* and *habit* to describe expressions: “it seems probable that some actions, which were at first performed consciously, have become through habit and association converted into reflex actions, and are now so firmly fixed and inherited, that they are performed, even when not of the least use” (Darwin 1872, p.39).

Similarly, what was observed during this research could be described as knowledge transmitted through ‘*non-posted communication’*[36] (NPC) which, as is the case with NVC or NVB, is not explicitly said but can be understood. This point is crucial because it creates an empirical issue for researchers working with online environments: they should recognise that, when merely extracting data from a website, they will not be able to grasp all the knowledge that is conveyed on it. Extracting online information only would be the equivalent of listening to a conversation without looking at the interlocutors. It is key to understand this because if some online interactions are unobservable, researchers can try to obtain them through other means (e.g. by talking to users) or, at the very least, by acknowledging the limitations of using online data.

A second development of the model presented in Chapter 2 relates to the acquisition of beliefs by an individual. It was thought that it was not possible to detect, at the individual level, if someone had acquired a belief or not. As described in Figure 2.5, it was supposed that the retention of beliefs would only be observed at the network level through the overall ranking displayed on the website. However, it was also possible to detect retention, at both levels, through the actions of users. Concretely, as discussed in Chapter 6, users from each group determined what ‘acceptable’ content was. Participants were given very little guidance on which elements made the information they posted of adequate quality. Yet, thanks to the feedback given by the ratings of others and their

[36] This is a new term devised by the researcher, therefore it will not be found in existing literature.

overall ranking on the website, they modified their content to what was considered appropriate, leading to emergent group norms. For instance, as one user reflected, she realised that posts which contained a video got higher ratings, and this motivated her to include audiovisuals in her subsequent posts. Therefore, with this example, the inclusion of videos was an action that was observed both at the individual and at the network levels.

The previous case could also be linked to NVC and NVB because individuals did not write a comment that stated 'you must include a video to get higher ratings'. What is more, it could have happened that people giving the ratings may not have consciously given higher ratings to posts that contained audio-visuals, and yet some users adopted it as an 'unspoken rule'. Thus, beliefs were also seen as affecting users' actions, and these actions could be observed both at the individual and network levels. For this reason, a third definition of knowledge that should be included in this thesis is: "information capable of affecting individuals' behaviour, that they acquire from other members of their species through teaching, imitation, and other forms of social transmission" (Richerson & Boyd 2005, p.5)[37] [38].

Finally, the last point that should be raised is the importance of the website designs in the transmission of knowledge, and the effect these have on interpretation. When communication is mainly shared through written language, which is the case in the majority of SNSs and CCs, interpretation happens through sense-reading and sense-giving (Polanyi 1967). Thus, the example of the dichotomous scale highlights the importance of having an appropriate design because, if users feel the available ones restrict or cause others to misinterpret their opinions, they might opt for not using the scale. This creates an issue for researchers since some communication becomes non-observable, but mainly it increases complexity for users, as they need to understand what a 'non-rating' means. In evolutionary terms, this means that the designs of websites have an impact on both the interactor and replicator. Regarding the interactor, the different set-ups constrain what users can post and the manner in which they give feedback to others. Consequently, this has an impact on the way they can express their beliefs (i.e. sense-

[37] Note that the definition was originally given to define 'culture', instead of 'knowledge'.
[38] Other scholars studying evolution have adopted similar definitions of culture, based partly on the one outlined above; for instance: "information that is acquired from other individuals via social transmission mechanisms such as imitation, teaching, or language" (Mesoudi 2011, pp.2–3).

giving), and also on how other users interpret them (i.e. sense-reading). Therefore, poor website designs can translate into higher interpretation issues or less engagement by users; as was the case with participants in the study who experienced the dichotomous scale.

8.2.2 Conceptualising the VSR process at the network level

The first development at the network level has already been mentioned. As previously outlined, the retention mechanism can also be observed if users adopt or change their behaviour. Behaviours can be described as a person's 'actions' (Darwin 1872) when analysed at the individual level. However, at the network level, *group behaviour* might better reflect this: "shared or collective reactions to others, systematically related to one's own and others' group memberships" (Turner 1984, p.522). As mentioned in Chapter 2, people vary their actions depending on their audience (Goffman 1959) and the group of which they are members (Tajfel & Turner 1979). In the same manner, after conducting the quasi-experiments, it was found that different cohorts adopted different behaviours, sometimes because individuals were different and sometimes because of the design of the website. For example, regarding behaviour caused by people's preferences, it was seen that different groups considered specific content more relevant and therefore it was posted at higher rates. Specifically, as mentioned in Chapter 6, one of the groups regarded videos as valuable, whereas another preferred case studies. Moreover, concerning the website's design, the dichotomous scale is a good example of how a whole group of individuals decided not to use the 'dislike' button as it was considered 'harsh'. Explained in Chapter 5, this was particularly noticeable for people who had a relationship in real-life, due to the avoidance of 'embarrassments' (Goffman 1959).

Second, as previously conceptualised, selection was observable only through the ratings given by users within the network. However, from the experiment, it became apparent that selection would involve at least one previous step that is also recorded on websites: accessing (i.e. clicking) the information. This might sound obvious, but it is useful to realise that online users make a 'pre-selection' of information and then go on to access and rate it. Researchers studying innovation adoption have described this as a two-stage process, which involves an awareness phase proceeded by a stage comprising evaluation and adoption (Van den Bulte & Lilien 2001). This is important because it may highlight dynamics within the network that would lead to a greater understanding of factors affecting selection. For instance, when comparing anonymous with identifiable users in Chapter 4, it was noted that not only did the ratings among different nationalities change but also the percentage of information accessed among distinct ethnicities. In this particular case, analysing the information accessed by users highlighted that they discriminate even before clicking on the information of a peer with a different nationality. The issue of nationalities will be further explained in the following sub-sections, yet it is important to note that both of these phases within the selection mechanism are worth analysing.

The third enhancement is that the VSR process was thought of as being 'linear', and with few interactions among factors. That is, it was thought that variation would affect selection and this would, in turn, affect retention. However, this was far from the truth: as observed during the quasi-experiment, the VSR process in online environments has more points of interaction between mechanisms. To begin with, most SNSs and CCs display the overall rankings in a very explicit way (e.g. by showing rating averages, counts of likes, dislikes, or helpful votes). When a new user enters the site, it takes very little time to identify the most relevant content. Therefore, the network's retained information would have a direct impact on the information users access. Moreover, by displaying other users' ratings, individuals will tend to use the aggregated values as a starting point for their own choices. Further, people who generate content get feedback in two ways: first by each rating that other users give, and then by the overall retention from the network, where they can see how their content is ranked compared to that of others. Therefore, the instant feedback that is characteristic of online environments makes selection and retention influence variation, and this happens at a higher speed than in an offline context. Figure 8.4 below captures the distinct interactions that take place among the three evolutionary mechanisms when the transmission of knowledge takes place in online environments.

8.2.3 Knowledge evolution in social media: an enhanced model

This discussion leads to a development in how knowledge evolution was theorised in online environments, as illustrated through an enhanced model in Figure 8.4, with both individual and network levels of the process further explained below.

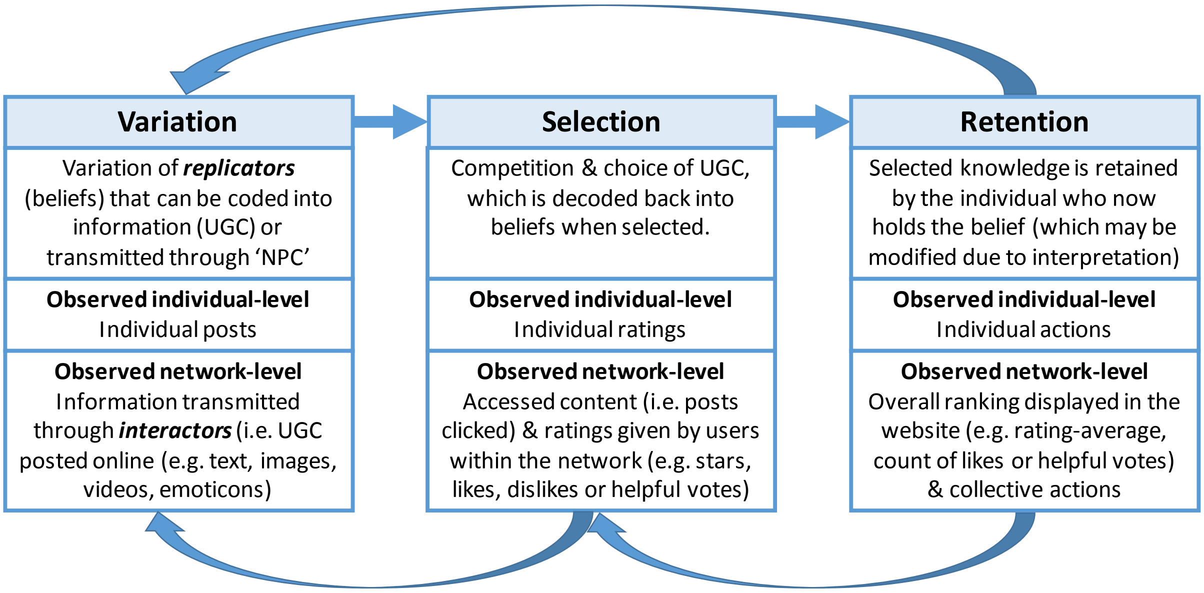


Figure 8.4 - Enhanced VSR process in social media

Individual level: The *variation* in beliefs (i.e. replicators) occurs when people either make variations of previous beliefs or form new ones. These are expressed through the interactors (i.e. UGC) and, in some cases, through non-posted communication (NPC). Only the former has been deeply studied in this research. *Selection* happens when individual recipients choose UGC expressed by other senders. Thus, selection involves an interpretation process as the content is decoded into knowledge (i.e. a new belief). Finally, *retention* happens when the receiver of the chosen UGC retains the belief. However, due to interpretation biases, the new belief might differ from the original one.

Network level: *Variation* takes place when a number of users within the network express their beliefs through UGC posted online, in the form of text, videos, images, emoticons, among others. However, over time, the beliefs expressed by individuals might homogenise around a common understanding, as individuals imitate those from their in-group. Further, *selection* can be detected through the information that individuals access and their evaluations of content (i.e. ratings). Both of these elements seem to be affected by the presence of the individuals' in-group. Regarding this point, people seem to perceive fewer differences from those close to them, and therefore have a greater 'trust' in the latter's content because – arguably – they perceive lesser errors in interpretation if they know the person sharing the information. Arguably, this makes the evaluations of the UGC shared by their in-group higher and makes people more prone to acquiring beliefs from those known to them. Finally, with time, collective beliefs are *retained* within groups. These can be observed through the network's rankings or rating-averages and, in some cases, also through users' collective actions. When a new user enters an online network, it becomes relatively easy to detect which knowledge is most valued by the group.

In summary, the difference between the first proposed VSR process (see Figure 2.5) and the enhanced model (Figure 8.4) is that, firstly, it was found that not all beliefs were coded into UGC, and some knowledge was transmitted through 'non-posted communication' (NPC). Further, the evolving entity (i.e. belief) was found to be affecting people's behaviour. For this reason, actions were added both at the individual and network levels. Moreover, it was found that there were two visible stages of selection: accessing information, followed by its evaluation. Further, additional connectors among the mechanisms were outlined and described. To conclude, this section advances the theory of knowledge evolution by defining what evolves in online environments, explaining how the VSR process occurs, and by highlighting new interactions between the mechanisms of variation, selection and retention.

Nevertheless, it should be highlighted that, although the VSR mechanisms have been studied in an online context – which allows for different interaction rules and communication practices than real-life environments – the research has enhanced the understanding of knowledge evolution, online and offline. For instance, the way in which knowledge is theorised to evolve by beliefs being coded into information, and then selected and retained by others with transmission errors, can be applied to a number of scenarios. Likewise, realising that both variation and selection can be heavily influenced by retention at the network level can be of value for any organisational study. For instance, when new members enter an organisation, they will be exposed to written information and unspoken rules that would dictate some of the beliefs that they adopt, which may have consequences for their behaviour. Finally, finding that every group will decide what 'acceptable' content is, and thus adapt its variations and feedback to this notion, can have implications for fields such as innovation and idea-generation.

Conclusion

This chapter summarises the thesis' proposition that website designs can be compared to *frames* because they can affect the judgement of individuals at the moment of choosing information. The research focused on two elements of website designs: user profiles (i.e. from anonymous to identifiable) and rating scales (i.e. from likert to dichotomous). Different user profiles affect the levels of self-disclosure and have a significant impact on the ratings of users because, by being identifiable, both the intergroup and interpersonal dimensions of situations become relevant. When this happens, users tend to self-categorise with their ingroups and therefore acquire more beliefs from them. Moreover, rating scales influence the levels of self-disclosure, which affect the choices of users significantly. Specifically, by making the rating dichotomous, there is a constraint in the interactor-replicator that increases transmission errors in communication and possibly generates higher levels of stress among decision-makers ('raters').

Furthermore, the degree of self-presentation results in a change of the hierarchy of biases, while the level of self-disclosure affects the intensity of the group-biases. These findings provide further evidence to support the idea that individuals' rationality is bounded by elements available at the moment of making a choice. Finally, users show higher levels of conformity towards their personal networks, which are proportional to the strength of their ties, a phenomenon explained by a mixture of social identification with the group and avoidance of real-life embarrassments with friends. However, it was also noted that higher conformity could translate into a reduced variation of beliefs and 'less smart' choices by the 'crowd'. Following all these findings, it can be concluded that the way in which websites are designed acts like a frame to the UGC posted online. Hence, users presented with the same information on social media sites with different designs might well rate that same information differently.

Lastly, this thesis contributes by conceptualising the development of knowledge in online environments through an evolutionary lens; first, by theorising how the VSR mechanisms take place online; and second, by explaining the selection mechanism through choice-making, identity, groups and relationships. Based on this discussion, the following and final chapter outlines a number of theoretical and methodological contributions, followed by a discussion of the practical and policy implications, limitations of the study, future research, and overall concluding remarks.

CHAPTER 9: SUMMARY AND CONCLUSION

This chapter constitutes the final part of the thesis. The first chapters of this thesis introduced the key concepts and theories, and established the aim of the research. Namely, to compare how the different designs of SNSs and CCs might enable, affect, or restrain the transmission of UGC, through the impact that these designs had on identity, groups and relationships, and consequently on the choices of users. This led to the empirical examination of three quasi-experimental conditions comprising approximately 1,000 participants, 200,000 online interactions, 400 questionnaires, and 6 focus groups. This work has resulted in a number of contributions to theory, methods, practice and policy. These contributions, followed by the limitations and challenges for future research, are highlighted and integrated in this last chapter.

9.1 THEORETICAL CONTRIBUTION

9.1.1 Knowledge evolution in social media: the wider contribution

By conceptualising the process of variation, selection and retention of knowledge in social media and proposing to study selection through choice-making, this thesis contributes to the field of online networks by providing a comprehensive rationale of how UGC content is transmitted and how/why people make choices online.

Making reference to the building blocks of theory development (Dubin 1978; Whetten 1989), the present research addresses the '*what*' element by outlining which factors ought to be considered as part of the explanation of online choice-making. Specifically, the proposed model takes an important step towards identifying individual, group, and relationship factors that help explain the selection of UGC. In addition, regarding the '*how*' block, the thesis thoroughly explains how the outlined factors relate. That is, a theoretical contribution of this thesis is to explain how theories of self-presentation, self-disclosure, group biases, strength of ties, and judgement work together into explaining the studied phenomena. Also, the research not only relies on the known connections among factors but also identifies new relationships. Lastly, concerning the '*why*' block, this thesis has sought to explain the underlying psychological and social dynamics that justify the inclusion of factors and their proposed casual relationships. The result is a robust model that can be used in a number of fields (e.g. social networks, online reviews, electronic word-of-mouth, online learning and education, etc.) to predict, explain and justify the choices of online users. Figure 9.1 combines the conceptual model of this thesis (Figure 8.1) with the enhanced VSR process (Figure 8.4), and highlights this thesis' main contribution: the selection of factors and the explanation of the *hows* and *whys* of their casual relationships. Additionally, the diagram below highlights those relationships that could be predicted by existing literature (in green), and those that are completely new, arising from this research (in orange):

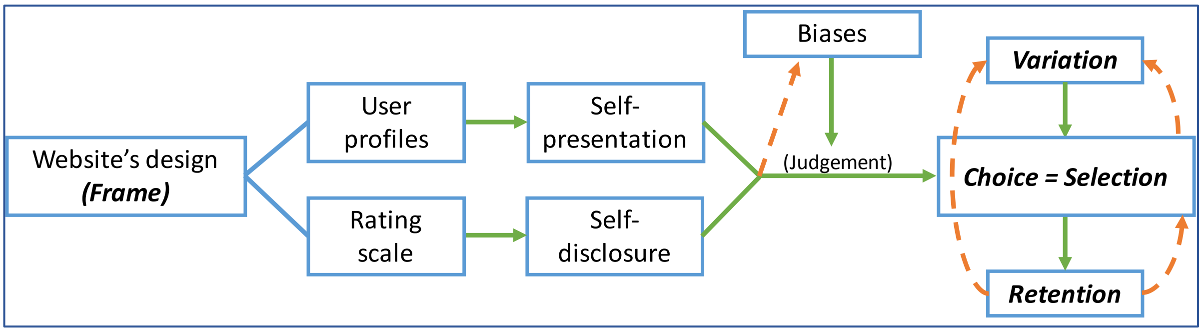


Figure 9.1 – Conceptual model with VSR

Nevertheless, the contribution of this work does not rely on the proposed model *per se*, but rather in its 'value-added contribution to theory development' (Whetten 1989, pp.492–494). The wider value-added contribution of this research is that it shows how the different frames adopted by social media websites – through particular elements of their designs – affect the transmission of UGC. Specifically, certain designs make the transmission of UGC more prone to biases and transmission errors. As a consequence, this affects not only the selection of content but also its variation and retention. In other words, certain elements of website designs affect the way in which users express and acquire beliefs and, subsequently, the knowledge that is created, chosen, and retained in the network. Therefore, by adopting a particular frame, websites can affect positively or negatively on the 'wisdom of the crowd'.

The consequences and implications of the above paragraph are vast, and this is what makes the present research so relevant and timely. On the one hand, the fast pace of social media and the huge amounts of users and data that it involves, makes it hard for researchers to analyse it. Proof of this is that, at the time of writing, the number of Facebook users comprises almost a third of the world's population; while, on average, 842 Instagram photos, 1,369 Tumblr posts, 8,014 Tweets, and 73,569 YouTube videos are posted online in just *one* second, worldwide (Internet Live Stats 2018). On the other hand, there seems to be little understanding of the many factors that affect the transmission of content, and it is hard to know what is causing – and how to solve – issues such as the spread of 'fake news'. Notably, with the recent scandal of *Cambridge Analytica*, researchers, journalists and online users are left wondering the extent to which social media influenced the outcomes of Brexit and various elections worldwide, including the one in the USA (Greenfield 2018). Users have "started to feel beat up by the same platforms and technologies that had enriched, empowered and connected" their lives (Friedman 2018, p.1), they resent trusting social networking sites (Wong 2018), and are becoming worried about the dominance of Facebook and other similar technology giants (Verkhivker 2018). This research has found that a way to begin tackling some of these issues is through improving the designs of websites, starting by how these setup user profiles and rating scales. This thesis has taken an important step towards identifying factors that can reduce biases and transmission errors in online environments, which can – therefore – have a positive impact on the creation, selection and retention of knowledge. Hopefully, if website designs improve and users become aware of particular aspects that can influence their choices, online 'crowds' can become wiser.

9.1.2 Specific contributions to research

Concretely, this thesis adds to other lines of research through having addressed the gaps in the literature outlined in Chapter 2 (see Table 2.4). Hence, this thesis contributes to the literature by:

- Advancing social media and platform design research by improving the existing knowledge about the diffusion of information in different types of sites/designs (e.g. Brandtzæg et al. 2009; Quan-Haase & Young 2010; Riedl et al. 2010; Hughes et al. 2012; Kietzmann et al. 2012; Riedl et al. 2013). This enhances the understanding of how certain aspects of a website's design – and different social media sites – affect the transmission of UGC.
- Extending evolutionary research (e.g. Campbell 1960; Weick 1969; Richerson & Boyd 2005; Mesoudi 2011) and the modest body of literature that uses the concept

of knowledge evolution to study online environments (Chen & Liang 2011; Kump et al. 2013; Jiang et al. 2014) by theorising for the first time how the VSR mechanisms take place in social media. Specifically, this thesis has advanced the theory of knowledge evolution by defining what evolves in online environments (i.e. the duality of beliefs-UGC) and by explaining how the VSR process occurs. Moreover, it was detected that there are more causal relationships between the three mechanisms than what the existing literature suggests (see Figure 8.4 and Figure 9.1), which can modify the way in which the VSR mechanisms are studied online and offline. Lastly, it was proposed and described that not all knowledge may be coded into information, which poses a number of challenges for scholars studying online environments.

- Adding to the body of research studying the effects of online anonymity (e.g. Lea & Spears 1991; Reicher et al. 1995; Postmes et al. 2001; Douglas & McGarty 2001; Lu & Bol 2007; Howard et al. 2010; Whittaker & Kowalski 2015). This work has increased the comprehension of how online users construct and manage their identities. Additionally, it has further contributed to literature by comparing how anonymous and *nonymous* users manage their identities (see below).
- Improving the current understanding of *nonimity* by further explaining how being identifiable online can affect real-life interactions and how/why users might avoid *embarrassments* (e.g. Zhao et al. 2008; Kietzmann et al. 2012).
- Extending the present knowledge of how the use of different rating scales affects the quality of ratings (Riedl et al. 2010; 2013). The existing knowledge was extended by studying rating scales within 'real' online settings, which allowed to unravel aspects such as conformity to the group in the context of two different rating scales.
- Advancing the understanding of knowledge transmission and group-biases (e.g. Durham 1991; Henrich 2001; Richerson & Boyd 2005; Mesoudi 2011) by applying these concepts to the study of online environments. A major advancement in this field was to discover a causal relationship between social processes and group biases that was non-existent in the literature. Namely, thanks to applying the theories of knowledge transmission and group biases to an environment that allows for different self-construction and management, it was

discovered and documented how self-presentation and self-disclosure affect the hierarchy and strength of group biases (see Figure 8.1 and Figure 9.1).

- Widening the study of conformity, which has received limited attention offline (Richerson & Boyd 2005) and, as shown in Chapter 2, has been ignored by online researchers (e.g. Riedl et al. 2010; Wu et al. 2011; Wilkinson & Thelwall 2012; Riedl et al. 2013; Liu & Park 2015; Park & Nicolau 2015). Most notably, this research argued for a differentiation between conformity to the whole and personal networks (i.e. outgroups and ingroups), consequently improving also the existing knowledge on *social approving cues* (e.g. Mueller et al. 2018).
- Addressing the literature gap concerning the effect of real-life tie strengths online (Kietzmann et al. 2012). This is probably one of the biggest contributions of this thesis given that, so far, no other study had investigated the impact of real-life friendships and strength of ties on online interactions. Hence, by making a further differentiation of ingroups (e.g. Tajfel & Turner 1979; Turner 1984; Turner & Reynolds 2012) using the theory of strength of ties (Granovetter 1973; 1983), this thesis explored – and reflected on – the effects of real-life relationship strengths on online interactions. Therefore, this research also contributes to the studies attempting to model real-life relationships with online interactions (e.g. Gilbert & Karahalios 2009).

However, "the route to good theory leads not through gaps in the literature but through an engagement with problems in the world that you find personally interesting" (Kilduff 2006, p.252). Thus, in addition to having fulfilled a substantial number of literature gaps, this thesis has addressed an issue that the researcher – and many journalists, managers, researchers, and society in general – find fascinating and worth-addressing. Social media is changing the rules of how people interact and disseminate information. It allows for identity constructions that individuals had never experienced and, most importantly, it has become the world's largest repository of human knowledge. Thus, if managed well, social media websites can improve the transmission of content among users, the choices people make online, and – ultimately – the evolution of knowledge.

To sum up, this thesis theoretically contributes to the understanding of how people make choices online, and how this affects the overall evolution of knowledge in social media. Researchers need to persevere in unravelling many of the psychological, sociological, and design factors that affect the variation, selection, and retention of UGC in online environments. This type of research is especially relevant in an ever-increasing user-generated web that billions of people use to communicate and acquire knowledge that influences many of their choices.

9.2 METHODOLOGICAL CONTRIBUTIONS

This thesis' quasi-experiment makes a methodological contribution through the way it was designed, the different sources from which data were collected, and the novel application of data analysis techniques such as SA in this context. First, regarding its design, this research has been unique in making use of an existing online (educational) community and performing changes to the way it is set up. These changes were inspired by the main characteristics of CCs and SNSs concerning self-presentation and self-disclosure, explained in Chapter 2. This comparison would have been impossible to perform if data had merely been extracted from an online environment, given that the type of media, content, and user demographics can vary substantially among different websites. Therefore, this study benefited from being real and malleable at the same time. Thanks to this, it has been possible to compare two distinct environments and determine the extent to which the set-up of a CC or a SNS restrains, affects, or enables the transmission of information.

Second, regarding data collection (i.e. types of data sources), it was argued in Chapter 3 that this is one of the most complete studies performed in social media, also due to the use of mixed- and multiple-methods. In fact, as outlined in Chapter 3, research that is based only on data extracted from SNSs or CCs has the limitation that all is known about users is what is available online. These studies, therefore, do not capture people's personal networks, nor are they able to grasp the 'whys' behind users accessing particular content or giving certain ratings (e.g. Wu et al. 2011; Wilkinson & Thelwall 2012; Procter et al. 2013; Thelwall et al. 2012; Cheng & Ho 2015; Park & Nicolau 2015; Liu & Park 2015; Goel & Goldstein 2014; Godes & Mayzlin 2004; Jacobsen 2015). On the other hand, controlled experiments have the advantage of allowing changes to the design of the online platform and securing in-depth input from participants. However, they miss some of the essential aspects of social media such as the influence of the network and, in particular, personal networks (e.g. Riedl et al. 2010; 2013; Mueller et al. 2018).

Through the use of mixed- and multiple-methods, this thesis aimed to overcome the limitations of single method approaches, in order to find robust answers to the research questions posed and literature gaps identified. As highlighted by other researchers, collecting data from a range of sources can help to get a better understanding of identity construction in different online environments (Zhao et al. 2008) and can provide more reliable results concerning collective decision-making (Riedl et al. 2010). For this reason, this study was based on virtual interactions, a survey that captured participants' personal networks and their impressions of the website's design, and focus groups where individuals were able to share more in-depth thoughts and feelings about the different set-ups and the effect of their personal networks on their ratings. Regarding the three sets of online interactions, each was naturally created over the course of a semester, and its comparison allowed the researcher to evaluate the distinct set-ups through the different rating patterns that emerged. Also, when the real-life personal networks of participants were added to the analysis, it was possible to determine their effect on ratings; and these results were independent of participant's interpretation. Furthermore, both questionnaires and focus groups served to obtain the perceptions of users regarding what had influenced their ratings, which enabled a comparison to be drawn between the evidence and people's interpretation of what took place. In addition, surveys and interviews also served to unravel the attitudes that users had towards different elements of the website's set-up, and to obtain a better understanding of how these and the presence of others affected the way in which users (students) accessed and evaluated UGC.

Third, concerning data processing, the study benefited from a wide range of analyses that varied between thematic – for qualitative data – to statistical, network, and sequential for quantitative data. As discussed in Chapter 3, this study made use of a post-positivist theoretical lens to address the research questions. Under post-positivism, researchers are encouraged to obtain multiple observations from different sources to triangulate across various imperfect perspectives (Trochim et al. 2016). To this end, the researcher made use of quantitative and qualitative data coming from the three sources described in the previous paragraphs. Triangulation is meant to help achieve convergent validity and completeness (Yu 2005) through the use of different methods of investigation, sources of data, and the cross-check of findings (Bryman & Bell 2011). However, the qualitative and quantitative data gathered was not only used for validation of findings and cross-checking. Instead, the overall analysis of the thesis became an iterative process that helped explain different parts of the issue of information transmission within online environments. That is, as repeatedly mentioned within the results chapters, online interactions guided some of the questions of the focus groups, while the responses obtained from these and the surveys gave rise to a number of quantitative analyses.

Nonetheless, the most significant contributions regarding the analysis of data have been the use of social network and sequence analyses. Regarding SNA (e.g. Granovetter 1973; 1983; Cross, Parker, et al. 2001; Cross, Borgatti, et al. 2001; Borgatti & Foster 2003), this research has advanced the study of the strength of ties in online environments by comparing four types of bonds – absent, weak, intermediate and strong – in the context of three website designs. Additionally, ties were not only inferred as in previous studies (e.g. Gilbert & Karahalios 2009); instead, this thesis also compared existing friendship ties with those inferred from 'followers' and drew conclusions from that. Further, the comments obtained from users through the surveys and focus groups illuminated some of the reasons why people might rate their personal networks in a certain way. Finally, including SNA was useful for more than simply differentiating personal networks and strengths of relationships. As seen in Chapter 6, SNA also allowed the researcher to unravel interactions within and between clusters of individuals (Table 6.4) and to map geodesic distances for all the individuals in the network (Table 6.5). To sum up, the present thesis has provided a wide range of applications of SNA, which should lead to further research using similar techniques that help obtain a better picture of the transmission of UGC within online networks, especially with the presence and influence of real-life relationships.

Finally, concerning SA (e.g. Gabadinho, Ritschard, Studer, et al. 2011; Gabadinho, Ritschard, Mueller, et al. 2011), one of the most significant contributions of this thesis has been to apply this analysis to the study of social media and ratings. As of May 2018, there were no published papers in which SA had been applied to social media, reviews, or ratings. Therefore, this study contributes by showing the many advantages that this form of analysis can have for scholars interested in interpreting (the sequence of successive states within) reviews, ratings, or idea evaluation. Moreover, for this particular research, Chapter 7 showed how SA helped to investigate the effects of conformity among the three quasi-experimental settings, thus providing the final evidence that different designs would generate different rating patterns, depending on the levels of conformity. No other online research to date has demonstrated the effect that user profiles and different rating scales can have on users. Based on this finding, a number of practical and policy implications are outlined and discussed in the following sections.

9.3 PRACTICAL IMPLICATIONS

Salient practical implications derive from the work advanced in this research. Table 9.1 presents a summary of the most relevant, which can be used by practitioners to make decisions regarding the design of user profiles and rating scales in online environments. It should be noted that is not intended to be prescriptive, but rather a tool for facilitating discussion of the different options available.

The table below presents a similar arrangement to the social media classification outlined by Kaplan & Haenlein (2010), which was adopted and described in Chapter 2 (see Table 2.1). However, Table 9.1 only varies the levels of self-presentation and self-disclosure. Self-presentation was modified by using different user profiles: anonymous and identifiable; self-disclosure was represented by the use of two rating scales: dichotomous and likert. Further, the content from the table summarises the outcomes from the quasi-experimental conditions regarding the prevalent group-bias and the levels of conformity obtained through the analysed transition rates. Once again, the content from Table 9.1 is not thought of as being prescriptive. However, it is believed that the type of bias and the strength of conformity follow a similar pattern in SNSs and CCs. Hence, this means that the websites that will produce the highest conformity are those that exploit users' personal networks and have narrow rating scales. In contrast, social media platforms that will evidence lower levels of conformity are those with anonymous users and broad scales. Furthermore, on sites where users are completely anonymous, there will be a tendency to rely more heavily on users' online-gained prestige. Conversely, platforms that show people's real identities will be more biased towards the similitude (i.e. homophily), relationship, or degrees of separation between the person authoring the content and others who are accessing and evaluating it.

Table 9.1 - Combined effects of self-presentation and self-disclosure

		USER PROFILES (→ SELF-PRESENTATION)	
		Anonymous (→ Low S-P) **Pros:** Users are more prone to giving real feedback. Ethnicity and gender do not play a role in ratings **Cons:** Users can behave competitively **Advice:** Use gamification to increase sharing of UGC and trust among users	**Identifiable (→ High S-P)** **Pros:** Users behave collaboratively **Cons:** People's 'personal front' (e.g. gender, nationality, etc.) becomes significant. Users evaluate UGC in line with their self-image and group-membership and not due to its quality **Advice:** Make use of existing personal networks to disseminate information
RATING SCALE (→ SELF-DISCLOSURE)	**Likert (→ High S-D)** **Pros:** Facilitates communication. Users are more prone to being critical with content **Cons:** Easier for people to feel criticised **Advice:** To facilitate communication even more, give users clear guidance of how to interpret values in the scale.	**Prevalent bias: Prestige** **Strength of conformity: Low** **Better for:** Educational websites and/or knowledge-sharing forums	**Prevalent bias: Conformity** **Strength of conformity: Medium** **Better for:** Environments where people should behave according to their 'real' selves, or when users are encouraged to rely on social ties (e.g. a company's internal social network)
	Dichotomous (→ Low S-D) **Pros:** Easy to distinguish when users like or dislike something. Makes users feel good. **Cons:** Only two options to communicate creates interpretation issues. People avoid disliking to the extent of rather not rating **Advice:** Use only in circumstances when the choices being taken are unimportant	**Prevalent bias: Prestige** **Strength of conformity: Medium** **Better for:** Could be suitable for websites where information is not meant to affect important choices and users are not going to be affected if others 'dislike' their posted UGC (e.g. a community dedicated to music/movies advice).	**Prevalent bias: Conformity** **Strength of conformity: High** **Better for:** Due the high levels of conformity observed, it is recommended to avoid this design for ethical reasons. Nevertheless, precisely due to high conformity, this is the ideal environment for marketers and politicians wanting to spread a belief at any cost.

Therefore, Table 9.1 has implications for practitioners who are designing online educational sites as well as for those setting up SNSs and CCs. A number of possible design guidelines can be given. Regarding online education and user profiles, anonymity led to lower levels of conformity and a lower rating-average, which could arguably mean that students were more analytical/critical. Further, as shown in Chapter 4, when students were anonymous, engagement was higher on average in terms of the number of questions, answers, ratings, comments, and badges per user. Thus, as being critical is seen as a desirable characteristic for students (Lu & Bol 2007; Howard et al. 2010; Li 2017), and given the focus among educators on student engagement (Trowler 2010; Kahu 2013), practitioners within online education should opt for an anonymous set-up. However, even when anonymous, it should not be forgotten that students are still influenced by the whole network regarding which content to access and how to evaluate it, and therefore it is also recommended that the individual and average ratings of previous users be hidden, at least until the current user has rated the question. Moreover, concerning rating scales, the best would be to adopt a likert scale, as it generates less conformity and because students believe that it does not unreasonably restrict their ability to assess, unlike the dichotomous scale. Nevertheless, as has been discussed, students using the likert scale also felt the evaluations they received were less fair, so this perhaps means that teachers and lecturers also need to prepare students to give and receive criticism.

Furthermore, although generalising outside educational environments – where this research took place – should be undertaken with caution, this study presents numerous insights into the broader field of social media. After all, the quasi-experimental conditions were designed based on the most characteristic elements of SNSs and CCs regarding user profiles and rating scales. Therefore, Table 9.1 shows how information transmission will be affected by these two types of social media. On the one hand, Table 9.1 could act as a guide for businesses that want to set up a website that makes use of ratings. In this case, companies could generate a mix of self-presentation and self-disclosure that adjusts to their goals. For instance, if someone wanted to create a forum with high levels of engagement, regardless of the comments of users being 'harsher', they might opt for an anonymous design. Contrarily, if they wanted a site where the comments of people are 'nicer', they might prefer to set it up with identifiable users. Further, companies seeking to advertise their products could use Table 9.1 to decide for which sites it is better to promote products through opinion leaders, and in which ones it is better to rely on people's personal networks. Hence, the results of this thesis could help CCs to evaluate the value of partnering with SNSs such as Facebook and Google+ be able to make use of their customers' personal networks. Lastly, regarding rating scales, sites wanting to use the 'wisdom of the crowds' (i.e. their customers' average ratings) to make predictions, could use Table 9.1 to decide when is better to use a likert or a dichotomous scale.

On the other hand, these results can also serve to question existing social media sites, and the changes they are implementing. For instance, as was explained in Chapter 2, during the last five years, many SNSs and CCs have been changing their rating scales, and seem to be doing this in a 'trial and error' manner. Also, many CCs have been merging with SNSs and seem to be exploiting users' personal networks, by highlighting not only their friends but even their friends-of-friends choices and preferences. What is more, SNSs are even automatically generating 'smart lists' based on similarity and declared friendships, and are therefore presenting users primarily with UGC from those with whom they share the same location, education, job, or social ties. Many changes are happening simultaneously and at a very high speed, and there seems to be very little consensus regarding the extent to which these sites might be influencing people's choices towards topics as banal as which tune to play next, to matters as important as for whom to vote in the next presidential elections.

Finally, quoting Riedl et al. (2010, p.16) "effective and accurate design of mechanisms for collective decision making is critical to harnessing the wisdom of the crowds. If the design is ill-fitted to the desired task, outcomes can be misleading or simply wrong". For this reason, the present thesis highlights the importance for practitioners of making informed decisions regarding the design of websites. However, as these stakeholders might have vested interests, it may be the job of policymakers to understand the effect that social media sites are having on the transmission of information, and to implement appropriate legislation towards their design.

9.4 POLICY IMPLICATIONS

Although the primary purpose of this research is not about policymaking, it is worth exploring some policy considerations based on what has been discussed so far. At the start of this thesis, social media was described as becoming a key communication channel used to interact with others, search for entertainment, read the news, and buy and sell products and services (Correa et al. 2010). However, in recent years social media has proven to play a more prominent role in our daily lives. To mention some examples, social media sites have shown to have a direct impact on businesses and politics (Fraser & Dutta 2008). Moreover, some websites and designs have been proven to produce addictive behaviours which arguably harm people because they reduce their attention span and increase their levels of anxiety (Soat 2015; Alter 2017). Further, ratings have been used in court as evidence for 'propagation of information', and a person was even convicted on defamation partly based on some of his Facebook 'likes' (DeVore 2017). Lastly, the most recent scandal from *Cambridge Analytica* suggests that social media might have been partly responsible for worldwide political and socio-economic decisions, such as *Brexit* in the UK, and the 2016 presidential elections in the US, together with over a hundred other election campaigns in over thirty countries (Ghoshal 2018; Greenfield 2018).

For these reasons, it is believed that there should be more research that helps policymakers understand the extent to which online environments might influence people's decisions. It is imperative for legislators to be better informed about these topics if they are to determine appropriate regulations for elements such as the design of websites. Unfortunately, most of the current research seems to focus on helping tech giants and businesses to implement ways of obtaining more data and engagement from users, and consequently to promote the marketing and diffusion of products and services (e.g. Gilbert & Karahalios 2009; Bonabeau 2009; Aral & Walker 2011; Kraut et al. 2011; Liu & Park 2015).

Therefore, it is hoped that this study can benefit legislators in understanding how social media might affect the transmission of information. This research has demonstrated that a significant percentage of online choices may be due to the actual set up of the platforms where we exchange information on a daily basis. And as this thesis measures only the effects of two of the four elements outlined by Kaplan & Haenlein (2010) – namely self-presentation and self-disclosure – the actual impact that the designs of SNSs and CCs have on the transmission of UGC is probably greater still.

Table 9.1 could assist policymakers in making certain key decisions regarding the design of websites. For instance, as outlined in Chapter 2 (Tables 2 and 3), in recent years many sites have been shrinking their rating scales, and most do not even allow for a 'thumbs down' button. Should this be allowed? Is it acceptable and safe for a whole generation of children and teenagers to grow up with sites that expect them to either 'like' or stay quiet? Is it good to condition people to only receiving positive feedback? Furthermore, regarding the presence of users' personal networks and the creation of automatic 'smart lists', how good is it for people to be exposed only to the opinions of those similar to them? As seen in Section 8.1, Richerson & Boyd (2005) noted that one of the main side effects of conformity is to reduce the amount of variation within groups, while maintaining the variation between groups. Thus, how good are these polarised echo-chambers for society? Is it fair for social media to have these effects on people? This thesis does not offer answers to these questions, as perhaps 'correct' responses do not exist; nonetheless, it is hoped the research offers potential insights for legislators and may contribute to the debate around such issues.

Finally, it should not be forgotten that "the adoption of a decision frame is an ethically significant act" (Tversky & Kahneman 1981, p.458). Thus, just as the food industry is required to follow specific rules and is asked to highlight when they use certain chemicals, online sites should perhaps be required to follow certain guidelines concerning their designs. Alternatively, at the very least, they might be required to inform users about the possible consequences of their use.

9.5 THESIS LIMITATIONS

As in any research, it is essential to acknowledge the limitations of the study, given that these may have influenced the findings of this thesis. The first limitation concerns the extent to which the findings may be considered generalisable. As explained in Chapter 2, various social media platforms were studied prior to setting up the experiment, with the intent to emulate some of the main characteristics of SNSs and CCs. Still, the site where the research was conducted had a particular purpose: education through peer interactions. Working with only one website allowed the findings to be attributed to the changes on its design, but also meant that some particular elements that are typical of video-sharing sites, music networks, or pure socialising platforms were not present. As a consequence, the results of this thesis might be strongly applicable to online communities that are similar to PeerWise but might be less relevant for sites which present more significant differences. To give an example, the findings of this research are probably very applicable to sites like TripAdvisor where users need to post reviews that are publicly accessible and are later up-voted by others within the site. Further, in that particular travel site, communication is done mainly by text, users obtain 'online-gained prestige' through gamification, and may sign-in with their 'real' identities or choose to remain anonymous. Conversely, the results of this thesis might be less applicable to *Snapchat*, where communication can only be done through pictures and videos that can be either public or shared with just a group of friends. Moreover, in this SNS which is based on existing social ties, the UGC is based on 'ephemeral messaging' and therefore disappears within a day of being posted. For this reason, it could not be claimed that this study is generalizable to all SNSs and CCs.

The second limitation concerns the sample. This research benefitted from having almost a thousand participants over the course of three years. Moreover, it had representatives of forty-nine nationalities based on two different countries. Nevertheless, individuals were all third-year undergraduates enrolled in the British education system. Thus, although their age group is the one with the most online users (Statista 2018c), not all internet users have the same levels of education which, arguably, is more likely to be accessed by people from particular socioeconomic groups. Therefore, a recommendation that will be further elaborated on in the following section is that future studies on the design of SNSs and CCs should consider including participants from other age groups, different levels of education and various socioeconomic statuses.

The third limitation has to do with missing data. To begin with, the number of students who responded to the questionnaire was 50-60% of those who took part in the first part of the data collection (i.e. online interactions). This percentage was obviously lower still with the focus groups. For this reason, the rationale behind some of the students' actions – which was mainly revealed from the qualitative data – might have not represented the whole population who took part in the study. Therefore, the observed online interactions may have additional explanations than the ones covered in this thesis. In addition, as explained in Chapter 3, the social network analysis (SNA) relied on a questionnaire, which required respondents to name “at least three but preferably five” personal connections from within the class. Therefore, because not every student replied to the survey, there was a loss of the personal connections between people. Moreover, some pupils probably had more than the three required acquaintances, and a few of them might have preferred not to name some of their friends. Hence, missing information on this question in particular is likely to have had an impact on the mapping of the network and its analysis. As a consequence, the SNA did not reflect or explain all the interactions, which might mean that the influence of users’ personal networks is probably higher than has been reported. Nonetheless, the attempt to map networks in this way is a unique endeavour in this context, and the data provide rich, significant and novel insights.

The last limitation concerns the possible self- consciousness of participants. As outlined in Chapter 3, the students from the selected modules were using PeerWise as part of an assignment. Hence, although nothing from the website counted directly towards their grades (e.g. getting right or wrong answers, receiving good or bad ratings or comments, badges or reputation scores), the fact that PeerWise was used for educational purposes might have made students behave differently than when they interact on an ‘unobserved’ platform. For instance, as argued by Zhao et al. (2008), people might not experience full disembodiment if they are using their institutional emails. Therefore, individual’s identity management might be different in SNSs and CCs. In particular, regarding ratings, some students declared to have been acting tactically to achieve better scores on the website (evidencing the power of gamification as a learning approach). However, it would perhaps be less likely for users to act in such a way on other social media sites, although users still care about number of followers, likes, retweets, and so forth. Finally, self-consciousness may have also been reflected in the self-report data collected with the survey and in the focus groups, leading participants to hide or emphasise issues that took place while interacting online. To mitigate these threats, the author collected data from multiple sources and participants were clearly advised that they could withdraw from the research at any time (Conway & Lance 2010). Nonetheless, it is recognised that, even then, people might behave differently in other online platforms when they are confident that they are not being ‘observed’.

9.6 FUTURE RESEARCH

The work presented in this research has provided a robust methodological approach and an initial set of results that can serve as the starting point for future research. Some of these paths of investigation can be drawn from the limitations of the work outlined above.

First, this research used one website to compare how different elements of the design of SNSs and CCs might affect the transmission of UGC. However, although this study has been instrumental in determining that different social media sites do enable or restrain the sharing of information, there is a limitation regarding its generalisation to specific sites. Thus, to address the generalisability issue outlined in the previous section, future research could focus on similar types of social media, instead of comparing among categories. For instance, prospective studies can focus on the study of music CCs or picture-based SNSs, and vary their designs to investigate if these changes affect the diffusion of content. Second, as previously mentioned, the sample population of this research was quite varied in a number of ways, but not regarding age groups, education and – to a significant extent – socioeconomic status. Therefore, apart from focusing on a diverse range of social media sites, future studies can also target different types of participants.

Third, regarding missing data, other studies could offer participants higher incentives for answering surveys and attending focus groups. The budget of the present study was modest given that it is for a doctoral research. However, higher incentives might increase participation, especially among students (Singer 2012). Fourth, some actions can be taken to reduce the self- consciousness of participants. To begin with, if future research were to be undertaken in the field of education, it might be worth doing so in a module where PeerWise – or any other educational site – is not utilised for assessment. Moreover, studies in social media could make use of publicly available sites, although they might need to find ways of ethically obtaining people's personal networks.

Fifth, regarding methodologies, as seen in Chapters 6 and 7, this study experimented with a range of analyses. Concerning networks, the most valuable outcomes came from geodesic distances. However, further research could deepen this analysis, exploring other SNA metrics such as different centrality measures, the connectedness of different clusters, degrees and weighted degrees of participants, and different modularity arrangements (e.g. Jackson 2008). Moreover, future research could also investigate those individuals who connect clusters and cliques, and also those relationships where both parties declared different tie strengths. That is, it may be insightful to have a deeper understanding of individuals who are responsible for connecting clusters that would otherwise end up as echo-chambers. Moreover, studying why and how different people have different perceptions of their relationship may be of use, especially if the way in which both accessed and evaluated each others' information can be compared. Furthermore, as outlined in Chapter 6, one analysis used was multilevel modelling (MM). These results hinted to be significant, which mean that it makes more sense to analyse data nested within clusters than with an overall regression. However, given that ratings involved repeated, bidirectional interactions among individuals nested in different clusters, a particular method within MM should be used, called 'multiple-membership, multiple-classification models' (Tranmer et al. 2014). This method was not pursued given that it currently does not allow for bidirectional analysis and it would have required much aggregation of the data, thus reducing the validity of results. However, it will be a future topic for research, which may even lead to enhance the current functionalities of the model.

Sixth, although the present research focused on UGC that was visible, an emerging finding was the existence of 'non-posted communication' (NPC, see Section 8.2.1). To date there appear to be only two papers dealing with tacit knowledge in the field of social media; both highlight the need for more research on this topic (Chatti et al. 2007; Panahi et al. 2012). However, neither of these papers studies more than what can be observed on the network, such as the effect of lack of comments or ratings. Therefore, NPC could be seen as a significant theoretical and empirical gap, and future research should tackle it.

Finally, future research could also study elements of the various underpinning frameworks that were omitted by design in this thesis. To begin with, from Campbell's (1960) outlined VSR mechanisms, this research focused on selection. Thus, future research could investigate in greater detail how the different designs of social media affect the variation and retention of knowledge. Moreover, as has been outlined from Chapter 2, this thesis focused on two of the four elements that generate the different types of social media: self-presentation and self-disclosure. However, the other two components outlined by Kaplan & Haenlein (2010), i.e. social presence and media richness, were not studied.

Thus, future studies could determine how changes in these two elements affect people's choices. Lastly, this thesis investigated how the different designs of social media might affect the transmission of information due to the effect that these would have on identity, groups and relationships. However, future research could also investigate the impact of websites' design on the other four blocks outlined by Kietzmann et al. (2011; 2012): reputation, conversations, sharing and presence. Specifically, the reputation block could be further developed with the existing datasets and a more in-depth analysis of the roles of badges and leaderboards in PeerWise. Namely, gamification could be studied in the context of the three quasi-experimental conditions.

As can be seen, there are numerous ideas for future research. Still, what is considered more important is that the present research might spark the broader interest of scholars, practitioners, and policymakers in understanding the importance of social media site design on the choices made by users.

9.7 SUMMARY AND CONCLUSION

As stated in Chapter 2, this research aimed to determine if – and how – different designs of social networking sites and content communities might enable, restrain or affect the transmission of user-generated content through the impact that these designs had on identity, groups and relationships. Based on a comprehensive review of the literature and after conducting rigorous empirical research that analysed the transmission of information through people's rating of UGC (i.e. their choices), the thesis' aim has been fulfilled and there has been a contribution to the body of research studying online networks. To sum up, this research proves and explains how two elements of website designs – user profiles and rating scales – act as a frame to the UGC that is posted online. Aditionally, they impact the transmission of information in two ways: 1) with errors in interpretation that cause original content to be misunderstood, and 2) by making some content more or less likely to be acquired, by affecting the type and strength of group-biases.

In terms of contributions, this research adds to theory by conceptualising the process of variation, selection and retention of knowledge in social media and proposing to study selection through choice-making. Additionally, it puts forward a model that studies the impact of frames (i.e. website designs) on choices (i.e. ratings), while making use of evolutionary, choice-making, identity, groups and network theories. Furthermore, regarding research on online environments, it makes a case for the distinction of whole and personal networks, in addition to online-gained and real-life prestige. Moreover, it advances the study of ingroups by making a further differentiation using the strength of ties, which also improves the knowledge of real-life relationships and their impact in online interactions. Lastly, the research adds to the body of literature on group-biases and conformity by discovering and documenting a link between the degrees of self-presentation and self-disclosure and the hierarchy and strength of biases.

Regarding methodology, the research contributes by conducting one of the most complete studies ever performed in online environments, thanks to the richness of the collected data and the access to participants' insights. By designing three quasi-experimental conditions that allowed the collection of data from a real online platform with actual users which had existent personal networks outside the network, this research benefited from real-life dynamics while still being close enough to users so that they could provide qualitative inputs. Further, regarding data analysis, the research presented an in-depth investigation of the effect of people's personal networks and the strength of ties, with a range of methods that went from simple rating averages to clustering analysis combined with geodesic paths. Finally, sequence analysis was applied for the first time to the study of social media and ratings, which provided insightful results that reflected the diverse degrees of conformity caused by different website designs.

Finally, the thesis provides a number of recommendations. Regarding practice, one of the research outcomes is a schematic that shows, for different websites designs, the most prominent bias and the 'levels of conformity' within online users. This table can provide practitioners within online education and those in charge of the design of social media with elements to decide what the 'ideal' design of sites should be; or at least, which aspects to avoid. Further, concerning policy, the study reflects on what it means for social media websites to adopt a frame, which is 'an ethically significant act' (Tversky & Kahneman 1981). Then, instead of providing detailed policy regulations, the study outlines a series of questions that could serve as the basis of debate among policymakers. Finally, the research concludes with an acknowledgement of limitations, followed by various possible lines for future research.

To conclude, it is believed that this study has offered a number of theoretical, empirical, practice and policy contributions, and it is foreseen that it will add value to fields such as social networks, online reviews, electronic word-of-mouth, online learning and education, among others. What is more relevant, is that this research can inform a better understanding of the impact that different designs have on choices, so practitioners and policymakers can implement website designs that allow for fruitful communication among people. By 2010 there were fewer than a billion people making use of social media, and this number is forecast to triple by 2021 when 3.02 billion are expected to have a SNS account (Statista 2018b). Thus, at current growth rates, it becomes imperative for the academic community, website creators, and governments to work together in ensuring that these platforms – which have become the biggest repositories of human knowledge and where people are accessing news, buying products and services, and discussing politics – have adequate designs that minimise biases and transmission errors. Improving website designs can positively influence the evolution of knowledge, aiding the ‘crowds’ to be wiser.

REFERENCES

Aldrich, H. & Ruef, M., 2006. *Organizations Evolving* 2nd ed., London: SAGE Publications.

Alter, A., 2017. *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked*, Penguin Press.

Altheide, D.L., 2000. Identity and the Definition of the Situation in a Mass-Mediated Context. *Symbolic Interaction*, 23(1), pp.1–27.

Amichai-Hamburger, Y. & Vinitzky, G., 2010. Social network use and personality. *Computers in Human Behavior*, 26(6), pp.1289–1295.

Andrade, E.B. & Ariely, D., 2009. The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes*, 109(1), pp.1–8.

Aral, S. & Walker, D., 2011. Creating social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science*, 57(9), pp.1623–1639.

Argyle, M., 1969. *Social interaction*, London: Methuen.

Argyle, M., Alkema, F. & Gilmour, R., 1971. The communication of friendly and hostile attitudes by verbal and non-verbal signals. *European Journal of Social Psychology*, 1(3), pp.385–402.

Ariely, D., Loewenstein, G. & Prelec, D., 2006. Tom Sawyer and the construction of value. *Journal of Economic Behavior and Organization*, 60(1), pp.1–10.

Ariely, D. & Zakay, D., 2001. A timely account of the role of duration in decision making. , 108, pp.187–207.

Asch, S.E., 1956. Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs: General and Applied*, 70(9), pp.1–70.

Bargh, J.A., McKenna, K.Y.A. & Fitzsimons, G.M., 2002. Can You See the Real Me? Activation and Expression of the “True Self” on the Internet. *Journal of Social Issues*, 58(1), pp.33–48.

Baron, R.S., Vandello, J.A. & Brunsman, B., 1996. The Forgotten Variable in Conformity Research: Impact of Task Importance on Social Influence. *Journal of Personality and Social Psychology*, 71(5), pp.915–927.

Barthelme, F., Ermine, J.-L. & Rosenthal-Sabroux, C., 1998. An architecture for knowledge evolution in organisations. *European Journal of Operational Research*, 109, pp.414–427.

Bastian, M., Heymann, S. & Jacomy, M., 2009. Gephi : An Open Source Software for Exploring and Manipulating Networks. In *Third International AAAI Conference on Weblogs and Social Media.* AAAI Publications.

Beal, V., 2016. Webopedia: Internet meme. Available at: http://www.webopedia.com/TERM/I/internet_meme.html [Accessed March 14, 2016].

Berger, J. & Iyengar, R., 2013. Communication Channels and Word of Mouth: How the Medium Shapes the Message. *Journal of Consumer Research*, 40(3), pp.567–579.

Bettman, J.R., Luce, M.F. & Payne, J.W., 1998. Constructive consumer choice processes. *Journal of Consumer Research*, 25(3), pp.187–217.

Bhalla, D., 2015. Standardized versus Unstandardized Regression Coefficient. *Listen Data*. Available at: https://www.listendata.com/2015/04/standardized-vs-unstandardized.html [Accessed September 3, 2017].

Blondel, V.D. et al., 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10).

Le Bon, G., 1896. *The crowd: a study of the popular mind*, London: T. Fisher Unwin, 1896.

Bonabeau, E., 2009. Decisions 2.0: The Power of Collective Intelligence. *MIT Sloan Management Review*, 50(2), pp.44–53.

Borgatti, S.P. & Foster, P.C., 2003. The network paradigm in organizational research: A review and typology. *Journal of Management*, 29(6), pp.991–1013.

Boudreau, M.-C., Gefen, D. & Straub, D.W., 2001. VALIDATION IN INFORMATION SYSTEMS RESEARCH: A STATE-OF-THE-ART ASSESSMENT. *MIS Quarterly*, 25(1), pp.1–16.

Boyd, R. & Richerson, P.J., 1985. *Culture and the Evolutionary Process*, University of Chicago Press.

Brandtzæg, P.B. et al., 2009. Norwegian Children’s Experiences of Cyberbullying When Using Different Technological Platforms. *Journal of Children and Media*, 3(4), pp.349–365.

Braun, V. & Clarke, V., 2008. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp.77–101.

Brendgen, M. et al., 2005. Examining Genetic and Environmental Effects on Social

Aggression : A Study of 6-Year-Old Twins. *Society for Research in Child Development,*, 76(4), pp.930–946.

Breslin, D., 2010. Generalising Darwinism to study socio-cultural change. *International Journal of Sociology and Social Policy*, 30(7/8), pp.427–439.

Breslin, D., 2011. Reviewing a Generalized Darwinist Approach to Studying Socio-economic Change. *International Journal of Management Reviews*, 13, pp.218–235.

Breslin, D., 2016. What evolves in organizational co-evolution? *Journal of Management & Governance*, 20(1), pp.45–67.

Breslin, D. & Jones, C., 2014. Developing an evolutionary/ecological approach in enterprise education. *The International Journal of Management Education*, 12(3), pp.433–444.

Bryman, A., 2012. *Social Research Methods* illustrate., OUP Oxford.

Bryman, A. & Bell, E., 2011. *Business Research Methods* 3rd editio., Oxford: Oxford University Press.

Van den Bulte, C. & Lilien, G.L., 2001. Two-Stage Partial Observability Models of Innovation Adoption. *Working Paper*, 3603(814), pp.1–45.

Burrell, G. & Morgan, G., 1979. *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*, Heinemann Educational.

Buxton, R., 2008. *Statistics: Multilevel modelling*,

Campbell, D.T., 1960. Blind Variation and Selective Retention in Creative Thought as in other Knowledge Processes. *Psychological Review*, 67(6), pp.380–400.

Campbell, D.T. & Ross, H.L., 1968. The Connecticut Crackdown on Speeding: Time-Series Data in Quasi-Experimental Analysis. *Law & Society Review*, 3(1), pp.33–54.

Cavalli-Sforza, L.L. & Feldman, M.W., 1981. *Cultural Transmission and Evolution: A Quantitative Approach* illustrate., Princeton University Press.

Chatti, M. a et al., 2007. The Web 2.0 Driven SECI Model Based Learning Process. In *Seventh IEEE International Conference on Advanced Learning Technologies, 2007 (ICALT 2007)*. pp. 780–783.

Chen, D.-N. & Liang, T.-P., 2011. Knowledge evolution strategies and organizational performance: A strategic fit analysis. *Electronic Commerce Research and Applications*, 10(1), pp.75–84.

Cheng, Y. & Ho, H., 2015. Social influence’s impact on reader perceptions of online reviews. *Journal of Business Research*, 68(4), pp.883–887.

Cheung, C.M.K. & Thadani, D.R., 2012. The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems*, 54(1), pp.461–470.

Ciancutti, J., 2011. Is there a better alternative from the 5-star rating system? *Quora*. Available at: https://www.quora.com/Is-there-a-better-alternative-to-the-5-star-rating-system [Accessed March 13, 2017].

Cloak, F.T., 1975. Is a cultural ethology possible? *Human Ecology*, 3(3), pp.161–182.

Colmer, R., 2013. Readability and the Coleman-Liau Index. *Readable*. Available at: https://readable.io/blog/the-coleman-liau-index/ [Accessed June 2, 2018].

Conway, J.M. & Lance, C.E., 2010. What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, 25(3), pp.325–334.

Correa, T., Hinsley, A.W. & de Zúñiga, H.G., 2010. Who interacts on the Web?: The intersection of users' personality and social media use. *Computers in Human Behavior*, 26(2), pp.247–253.

Cross, R., Parker, A., et al., 2001. Knowing what we know: Supporting knowledge creation and sharing in social networks. *Organizational Dynamics*, 30(2), pp.100–120.

Cross, R., Borgatti, S.P. & Parker, A., 2001. Beyond answers: Dimensions of the advice network. *Social Networks*, 23(3), pp.215–235.

Cross, R., Borgatti, S.P. & Parker, A., 2002. Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration. *California Management Review*, 44(2), pp.25–47.

Cunha, E. et al., 2011. Analyzing the Dynamic Evolution of Hashtags on Twitter: a Language-Based Approach. In *Proceedings of the Workshop on Language in Social Media LSM 2011*. Portland, pp. 58–65.

Cyr et al., 2009. Exploring Human Images in Website Design: A Multi-Method Approach. *MIS Quarterly*, 33(3), p.539.

Dabbagh, N. & Kitsantas, A., 2012. Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *Internet and Higher Education*, 15(1), pp.3–8.

Darwin, C.R., 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life* 1st ed. J. Van Wyhe, ed., London: John Murray.

Darwin, C.R., 1872. *The expression of the emotions in man and animals* 1st ed. J. Van Wyhe, ed., London: John Murray.

Dawkins, R., 1976. *The Selfish Gene* 30th Anniv., Oxford: Oxford University Press.

Dean, N., 2016. *Cluster Analysis using R*, Glasgow.

Dellarocas, C., 2003. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science*, 49(10), pp.1407–1424.

Denny, P., 2013. The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. pp. 763–772.

Denzin, N.K. & Lincoln, Y.S., 2018. *The Sage Handbook of Qualitative Research* Fifth edit., SAGE.

DeVore, V., 2017. "Liking" a Facebook post is tricky legal territory. *SWI SwissInfo*. Available at: https://www.swissinfo.ch/eng/society/liking-a-facebook-post-is-tricky-legal-territory/43228744 [Accessed January 8, 2018].

Distin, K., 2011. *Cultural Evolution*, Cambridge: Cambridge University Press.

Distin, K., 2006. Fame! Money! Power! *Think*, 5(13), pp.89–93.

Dobson, S. et al., 2013. Small firm survival and innovation: An evolutionary approach. *Entrepreneurship and Innovation*, 14(2), pp.69–80.

Donaldson, L., 1997. A positivist alternative to the structure-action approach. *Organization studies*, 18(1), pp.77–92.

Douglas, K.M. & McGarty, C., 2001. Identifability and self-presentation : Computer-mediated communication and intergroup interaction. *British Journal of Social Psychology*, (40), pp.399–416.

Dubin, R., 1978. *Theory building* 2nd revise., Free Press.

Durham, W.H., 1991. *Coevolution: Genes, Culture, and Human Diversity*, Stanford University Press.

Eisenhardt, K., 1989. Building Theories from Case Research. *The Academy of Management Review*, 14(4), pp.532–550.

Eryarsoy, E. & Piramuthu, S., 2014. Experimental evaluation of sequential bias in online customer reviews. *Information & Management*, 51(8), pp.964–971.

Fereday, J. & Muir-Cochrane, E., 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, 5(1), pp.80–92.

Field, A., 2013. *Discovering Statistics using IBM SPSS Statistics* 4th Editio., London:

SAGE.

Forgas, J.P., 1985. *Interpersonal Behaviour: The Psychology of Social Interaction* J. P. Forgas, ed., Pergamon Press.

Fraser, M. & Dutta, S., 2008. *Throwing Sheep in the Boardroom: How Online Social Networking Will Transform Your Life, Work and World*, Wiley.

Friedman, M., 1953. *Essays in positive economics.*, Chicago U.P. Cambridge U.P, 1953.

Friedman, T.L., 2018. How Mark Zuckerberg Can Save Facebook - and Us. *The New York Times*.

Gabadinho, A., Ritschard, G., Mueller, N.S., et al., 2011. Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), pp.1–37.

Gabadinho, A., Ritschard, G., Studer, M., et al., 2011. *Mining sequence data in R with the TraMineR package : A user's guide for version 1.8*,

Gephi, 2017. Gephi - The Open Graph Viz Platform. Available at: https://gephi.org/ [Accessed July 2, 2018].

Ghoshal, D., 2018. Cambridge Analytica scandal: All the countries where SCL Elections claims to have worked — Quartz. *Quartz*.

Gigerenzer, G., 2000. Bounded rationality: The adaptive toolbox. *International Journal of Psychology*, 35(3–4), pp.203–204.

Gilbert, E. & Karahalios, K., 2009. Predicting tie strength with social media. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. pp. 211–220.

Gill, J. & Johnson, P., 2010. *Research methods for managers* 4th ed., Los Angeles: SAGE.

Gilovich, T., Griffin, D. & Kahneman, D. eds., 2002. *Heuristics and biases: The psychology of intuitive judgment*, Cambridge University Press.

Godes, D. & Mayzlin, D., 2004. Using Online Conversations to Study Word-of-Mouth Communication. *Marketing Science*, 23(4), pp.545–560.

Goel, S. & Goldstein, D., 2014. Predicting Individual Behavior with Social Networks. *Marketing Science*, 33(1), pp.82–93.

Goffman, E., 1963. *Behavior in public places: Notes on the social organization of gatherings*, New York: The Free Press.

Goffman, E., 1959. *The presentation of self in everyday life*, London: Penguin-1990.

Gopinath, S., Thomas, J.S. & Krishnamurthi, L., 2014. Investigating the Relationship Between the Content of Online Word of Mouth, Advertising, and Brand

Performance. *Marketing Science*, 33(2), pp.241–258.

Granovetter, M., 1983. The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, pp.201–233.

Granovetter, M., 1973. The Strength of Weak Ties. *American Journal of Sociology*, 78(6), p.1360.

Greenfield, P., 2018. The Cambridge Analytica files: the story so far. *The Guardian*.

Guba, E.G. & Lincoln, Y.S., 1994. Competing Paradigms in Qualitative Research. In K. Densin & Y. S. Lincoln, eds. *Handbook of qualitative research*. California: Thousand Oaks SAGE, pp. 105–117.

Guest, G., MacQueen, K. & Namey, E., 2012. Applied Thematic Analysis.

Gutiérrez-Rubí, A., 2014. 6 rasgos clave de los millennials, los nuevos consumidores. *Forbes Mexico*.

Handelsman, J. & Moss-Racusin, C., 2012. Scientists not immune from gender bias. *YaleNews*, p.2. Available at: https://news.yale.edu/2012/09/24/scientists-not-immune-gender-bias-yale-study-shows [Accessed August 14, 2018].

Hastie, R. & Dawes, R.M., 2010. *Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making* 2nd ed. R. Hastie & R. M. Dawes, eds., California: SAGE.

Heath, C., Bell, C. & Sternberg, E., 2001. Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, 81(6), pp.1028–1041.

Hecht, M.A. & Ambady, N., 1999. Nonverbal communication and psychology: Past and future. *New Jersey Journal of Communication*, 7(2), pp.156–170.

Henrich, J., 2004. Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*, 53, pp.3–35.

Henrich, J., 2001. Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioural change. *American Anthropologist*, 103(4), pp.992–1013.

Henrich, J. & Boyd, R., 1998. The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, 19, pp.215–241.

Henrich, J. & Gil-White, F.J., 2001. The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22(3), pp.165–196.

Hill, S. & Ready-Campbell, N., 2011. Expert Stock Picker: The Wisdom of (Experts in)

Crowds. *International Journal of Electronic Commerce*, 15(3), pp.73–102.

Hodgson, G.M., 2005. Generalizing Darwinism to Social Evolution: Some Early Attempts. *Journal of Economic Issues*, 39(4), pp.899–914.

Howard, C.D., Barrett, A.F. & Frick, T.W., 2010. Anonymity to Promote Peer Feedback: Pre-Service Teachers' Comments in Asynchronous Computer-Mediated Communication. *Journal of Educational Computing Research*, 43(1), pp.89–112.

Hu, Y., Manikonda, L. & Kambhampati, S., 2014. What We Instagram: A First Analysis of Instagram Photo Content and User Types. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, pp. 595–598.

Huang, Y.Y. & Chou, C., 2010. An analysis of multiple factors of cyberbullying among junior high school students in Taiwan. *Computers in Human Behavior*, 26(6), pp.1581–1590.

Hughes, D.J. et al., 2012. A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage. *Computers in Human Behavior*, 28(2), pp.561–569.

Internet Live Stats, 2018. Internet Users. Available at: http://www.internetlivestats.com/internet-users/ [Accessed January 9, 2017].

Iyengar, R., Van den Bulte, C. & Valente, T.W., 2011. Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science*, 30(2), pp.195–212.

Jackson, M.O., 2008. *Social and economic networks*, Princeton, N.J.: Princeton University Press.

Jacobsen, G.D., 2015. Consumers, experts, and online product evaluations: Evidence from the brewing industry. *Journal of Public Economics*, 126, pp.114–123.

Jalilvand, M.R., Esfahani, S.S. & Samiei, N., 2011. Electronic word-of-mouth: Challenges and opportunities. *Procedia Computer Science*, 3, pp.42–46.

Jiang, G. et al., 2014. Evolution of knowledge sharing behavior in social commerce: An agent-based computational approach. *Information Sciences*, 278, pp.250–266.

Jiang, G. et al., 2015. Impacts of knowledge on online brand success: an agent-based model for online market share enhancement. *European Journal of Operational Research*.

Johnson, B.R. & Onwuegbuzie, A.J., 2004. Mixed Methods: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), pp.14–26.

Johnson, P. & Duberley, J., 2000. Understanding Management Research.

Kahneman, D., 2003. A Perspective on Judgment and Choice: Mapping Bounded Rationality. *American Psychologist*, 58(9), pp.697–720.

Kahneman, D., 2012. *Thinking, Fast and Slow*, London: Penguin.

Kahneman, D., Slovic, P. & Tversky, A. eds., 1982. *Judgment under uncertainty: Heuristics and biases*, New York: New York.

Kahneman, D. & Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), pp.263–291.

Kahu, E.R., 2013. Framing student engagement in higher education. *Studies in Higher Education*, 38(5), pp.758–773.

Kaplan, A.M. & Haenlein, M., 2010. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), pp.59–68.

Kietzmann, J.H. et al., 2011. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), pp.241–251.

Kietzmann, J.H. et al., 2012. Unpacking the social media phenomenon: towards a research agenda. *Journal of Public Affairs*, 12(2), pp.109–119.

Kilduff, M., 2006. Editor’s comments: Publishing theory. *Academy of Management Review*, 31(2), pp.252–255.

Kilduff, M. & Brass, D.J., 2010. Organizational social network research: Core ideas and key debates. *Academy of Management Annals*, 4(1), pp.317–357.

Kozinets, R. V., 2002. The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities. *Journal of Marketing Research*, 39(1), pp.61–72.

Krauss, R.M., Chen, Y. & Chawla, P., 1996. Nonverbal Behavior and Nonverbal Communication: What do Conversational Hand Gestures Tell Us? *Advances in Experimental Social Psychology*, 28, pp.389–450.

Kraut, R.E. et al., 2011. *Building Successful Online Communities: Evidence-Based Social Design*, MIT Press.

Krueger, R.A., 2015. *Focus Groups : A Practical Guide for Applied Research* 5th editio. M. A. Casey, ed., Thousand Oaks, California : SAGE, 2015.

Krug, S., 2016. Reactions Now Available Globally. *Facebook Newsroom*. Available at: http://newsroom.fb.com/news/2016/02/reactions-now-available-globally/ [Accessed March 13, 2017].

Kump, B. et al., 2013. Tracing knowledge co-evolution in a realistic course setting: A wiki-based field experiment. *Computers and Education*, 69, pp.60–70.

Lantz-Andersson, A., Vigmo, S. & Bowen, R., 2013. Crossing boundaries in Facebook:

Students' framing of language learning activities as extended spaces. *International Journal of Computer-Supported Collaborative Learning*, 8(3), pp.293–312.

Lasswell, H.D., 1948. The structure and function of communication in society. In L. Bryson, ed. *The communication of ideas: A series of addresses*. New York: Harper, pp. 117–130.

Lea, M. & Spears, R., 1991. Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies*, 34(2), pp.283–301.

Lee, D., Kim, H.S. & Kim, J.K., 2012. The role of self-construal in consumers' electronic word of mouth (eWOM) in social networking sites: A social cognitive approach. *Computers in Human Behavior*, 28(3), pp.1054–1062.

Lee, Y.-C., 2014. Impacts of decision-making biases on eWOM retrust and risk-reducing strategies. *Computers in Human Behavior*, 40, pp.101–110.

Lengerich, E., 2016. *Bias, Confounding and Effect Modification*,

Li, L., 2017. The role of anonymity in peer assessment. *Assessment and Evaluation in Higher Education*, 42(4), pp.645–656.

Litvin, S.W., Goldsmith, R.E. & Pan, B., 2008. Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, 29(3), pp.458–468.

Liu, Z. & Park, S., 2015. What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, pp.140–151.

Loasby, B.J., 2002. The evolution of knowledge: beyond the biological model. *Research Policy*, 31, pp.1227–1239.

Loomer, J., 2012. Detailed History of Facebook Changes 2004-12 [Research]. Available at: https://www.jonloomer.com/2012/05/06/history-of-facebook-changes/ [Accessed April 12, 2018].

Lorenz, J. et al., 2011. How social influence can undermine the wisdom of crowd effect. In *Proceedings of the National Academy of Sciences*. pp. 9020–9025.

Lowry, R., 2001. Log-Linear Analysis for an AxBxC Contingency Table. *VassarStats*. Available at: http://vassarstats.net/abc.html [Accessed June 6, 2018].

Lu, R. & Bol, L., 2007. A Comparison of Anonymous Versus Identifiable e-Peer Review on College Student Writing Performance and the Extent of Critical Feedback. *Journal of Interactive Online Learning*, 6(2), pp.100–115.

Lund, A. & Lund, M., 2015. Spearman's Rank-Order Correlation. *Laerd Statistics*. Available at: https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-guide.php [Accessed June 3, 2018].

MacLachlan, R., 2014. The Human Condition. Available at: http://humancond.org/ [Accessed March 9, 2017].

Malone, T.W., Laubacher, R. & Dellarocas, C., 2009. Harnessing crowds : Mapping the genome of collective intelligence. *MIT Sloan School of Management*, 1, pp.1–20.

Malone, T.W., Laubacher, R. & Dellarocas, C., 2010. The collective intelligence genome. *IEEE Engineering Management Review*, 38(3), p.38.

McCracken, H., 2016. Jack Dorsey On The New Twitter: "We're Not A Social Network As People Think About It". *Fast Company*.

Meng, M. & Agarwal, R., 2007. Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information Systems Research*, 18(1), pp.42–67.

Mesoudi, A., 2011. *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*, Chicago: The University of Chicago Press.

Miller, K., 2005. Post-Positivist Perspectives on Theory Development. In *Communication Theories: Perspectives, Processes, and Contexts*. Boston: McGraw-Hill, p. 355.

Mishna, F., Saini, M. & Solomon, S., 2009. Ongoing and online: Children and youth's perceptions of cyber bullying. *Children and Youth Services Review*, 31(12), pp.1222–1228.

Moon, Y., 2000. Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research*, 26(4), pp.323–339.

Mueller, J. et al., 2018. Reframing the decision-makers' dilemma: Towards a social context model of creative idea recognition. *Academy of Management Journal*, 61(1), pp.94–110.

Von Neumann, J. & Morgenstern, O., 1957. *Theory of games and economic behavior* 2007–60th ed. O. Morgenstern, O. Morgenshtern, & O. (Oskar) Morgenshtern, eds., Princeton, N.J.: Princeton, N.J.

Nonaka, I., 1994. A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), pp.14–37.

OECD, 2007. *Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking*, Paris.

Orlikowski, W.J., 2002. Knowing in Practice: Enacting a Collective Capability in

Distributed Organizing. *Organization Science*, 13(3), pp.249–273.

Oxford Dictionaries, 2018. Definition of Markov Chain.

Panahi, S., Watson, J. & Partridge, H., 2012. Social media and tacit knowledge sharing: developing a conceptual model. In *World Academy of Science, Engineering and Technology (WASET)*. Paris, pp. 1095–1102.

Park, S. & Nicolau, J.L., 2015. Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, 50, pp.67–83.

PeerWise, 2015. Welcome to PeerWise. Available at: https://peerwise.cs.auckland.ac.nz/ [Accessed July 16, 2015].

Perez, S., 2010. Twitter is NOT a Social Network, Says Twitter Exec. *readwrite*.

Perreault, C., Moya, C. & Boyd, R., 2012. A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior*, 33(5), pp.449–459.

Peters, K. et al., 2013. Social media metrics - A Framework and Guidelines for Managing Social Media. *Journal of Interactive Marketing*, 27(4), pp.281–298.

Polanyi, M., 1968. Logic and Psychology. *American Psychologist*, 23(1), pp.27–43.

Polanyi, M., 1998. *Personal knowledge [electronic resource] : towards a post-critical philosophy* 6th ed., London: Taylor & Francis.

Polanyi, M., 1967. Sense-Giving and Sense-Reading. *Philosophy*, 42(162), pp.301–325.

Popper, K.R., 1959. *The Logic of Scientific Discovery* revised., London: Routledge Classics, 2005.

PopulationPyramid, 2018. Population Pyramids of the World from 1950 to 2100. Available at: https://www.populationpyramid.net/europe/ [Accessed January 10, 2017].

Postmes, T. et al., 2001. Social Influence in Computer-Mediated Communication : The Effects of Anonymity on Group Behavior Social Influence in Computer-Mediated Communication : The Effects of Anonymity on Group Behavior. *Personality and Social Psychology Bulletin*, 27(10), pp.1243–1254.

Procter, R., Vis, F. & Voss, A., 2013. Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology*, 16(3), pp.197–214.

Quan-Haase, a. & Young, a. L., 2010. Uses and Gratifications of Social Media: A Comparison of Facebook and Instant Messaging. *Bulletin of Science, Technology & Society*, 30(5), pp.350–361.

Reicher, S.D., Spears, R. & Postmes, T., 1995. A Social Identity Model of

Deindividuation Phenomena. *European Review of Social Psychology*, 6(1), pp.161–198.

Richerson, P.J. & Boyd, R., 2005. *Not by genes alone: how culture transformed human evolution*, Chicago: The University of Chicago Press.

Riedl, C. et al., 2010. Rating Scales for Collective Intelligence in Innovation Communities: Why Quick and Easy Decision Making Does Not Get it Right. *Thirty First International Conference on Information Systems*, pp.98–104.

Riedl, C. et al., 2013. The Effect of Rating Scales on Decision Quality and User Attitudes in Online Innovation Communities. *International Journal of Electronic Commerce*, 17(3), pp.7–36.

Robson, C., 2011. *Real world research: a resource for users of social research methods in applied settings* Third Edit., Wiley.

Robson, K. et al., 2016. Game on: Engaging customers and employees through gamification. *Business Horizons*, 59(1), pp.29–36.

Roman, D., 2009. Crowdsourcing and the question of expertise. *Communications of the ACM*, 52(12), pp.12–12.

Romero, D.M., Meeder, B. & Kleinberg, J., 2011. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the 20th international conference on World wide web - WWW 2011*. pp. 695–704.

Rosenberg, M., 1979. *Conceiving the self*, Basic Books.

Rubin, A., 2010. *Essential research methods for social work* 2nd ed. E. R. Babbie, ed., Belmont, Calif.: Belmont, Calif. : Brooks/Cole Cengage Learning, c2010.

Schau, H.J. & Gilly, M.C., 2003. We Are What We Post? Self-Presentation in Personal Web Space. *Journal of Consumer Research*, 30(3), pp.385–404.

Schwartz, B., 2009. *The Paradox of Choice: Why More Is Less* Revised Ed., Harper Collins.

Schweisberger, V., Billinson, J. & Chock, T.M., 2014. Facebook, the Third-Person Effect, and the Differential Impact Hypothesis. *Journal of Computer-Mediated Communication*, 19(3), pp.403–413.

Sedghi, A., 2014. Facebook: 10 years of social networking, in numbers. *The guardian UK*.

SEO Web Marketing, 2012. Google Plus History: Google's Answer to Social Networking. Available at: http://www.seowebmarketing.co.uk/google-plus-history/

[Accessed April 18, 2018].

Siegel, D. a., 2013. Social Networks and the Mass Media. *American Political Science Review*, 107(04), pp.786–805.

Simon, H.A., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1), pp.99–118.

Simon, H.A., 1979. Rational decision making in business organizations. *American Economic Review*, 69(4), pp.493–513.

Singer, E., 2012. *The Use and Effects of Incentives in Surveys*, Washington DC.

Smith, D., 2017. Why Netflix's "thumbs up, thumbs down" ratings system is fundamentally flawed. *Business Insider UK*.

Smith, E.A. & Senter, R.J., 1967. *Automated readability index.*, Ohio.

Soat, M., 2015. Social Media Triggers a Dopamine High. *American Marketing Association*. Available at: https://www.ama.org/publications/MarketingNews/Pages/feeding-the-addiction.aspx [Accessed April 27, 2018].

Statista, 2018a. Most famous social network sites worldwide as of January 2018. *Statista*. Available at: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [Accessed April 6, 2018].

Statista, 2018b. Number of social network users worldwide from 2010-2021 in billions. Available at: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ [Accessed May 15, 2018].

Statista, 2018c. Share of adults in the United States who use the internet in 2018, by age group. Available at: https://www.statista.com/statistics/266587/percentage-of-internet-users-by-age-groups-in-the-us/ [Accessed May 14, 2018].

Stephenson, G.M., 1981. Intergroup Bargaining and Negotiation. In J. C. Turner & H. Giles, eds. *Intergroup Behaviour*. Oxford: Basil Blackwell, p. 277.

Stone, B.K. et al., 2016. How to Identify the Most Important Predictor Variables in Regression Models. *The Minitab Blog*. Available at: http://blog.minitab.com/blog/adventures-in-statistics-2/how-to-identify-the-most-important-predictor-variables-in-regression-models [Accessed June 3, 2018].

Stride, C., 2016. Introduction to Multilevel Modelling Using SPSS. , pp.1–89.

Suler, J., 2004. The Online Disinhibition Effect. *CyberPsychology & Behavior*, 7(3), pp.321–326.

Suler, J.R., 2002. Identity Management in Cyberspace. *Journal of Applied Psychoanalytic*

Studies, 4(4), pp.455–459.

Sun, X. et al., 2015. Topological evolution of virtual social networks by modeling social activities. *Physica A: Statistical Mechanics and its Applications*, 433, pp.259–267.

Surowiecki, J., 2004. *The Wisdom of Crowds: Why the Many are Smarter Than the Few and how Collective Wisdom Shapes Business, Economies, Societies, and Nations*, Doubleday.

Tajfel, H., 1978. *Differentiation between social groups: studies in the social psychology of intergroup relations*, London: Academic Press.

Tajfel, H., 1974. Social identity and intergroup behaviour. *Social Science Information*, 13(2), pp.65–93.

Tajfel, H. & Turner, J., 1979. An Integrative Theory of Intergroup Conflict. In W. G. Austin & S. Worchel, eds. *The Social Psychology of Intergroup Relations*. Brooks/Cole Pub. Co., pp. 33–47.

Teehan, G., 2016. Reactions: Not everything in life is Likable. *Facebook Design*. Available at: https://medium.com/facebook-design/reactions-not-everything-in-life-is-likable-5c403de72a3f#.cie3bti4k [Accessed March 13, 2017].

The McCarthy Group, 2014. *Millennials: Trust & Attention Survey*, New York.

Thelwall, M., 2009. *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*, Morgan & Claypool.

Thelwall, M., Buckley, K. & Paltoglou, G., 2011. Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2), pp.406–418.

Thelwall, M., Sud, P. & Vis, F., 2012. Commenting on YouTube videos: From Guatemalan Rock to El Big Bang. *Journal of the American Society for Information Science and Technology*, 63(3), pp.616–629.

Thiétart, R.-A. ed., 2001. *Doing management research: A comprehensive guide*, London: SAGE.

Thompson, D., 2009. Ranking predictors in logistic regression. In *Midwest Sas Users Group - MWSUG 2009 Conference*.

Tourangeau, R., Rips, L.J. & Rasinski, K.A., 2000. *The psychology of survey response* Online-201., Cambridge, UK: Cambridge University Press.

Tranmer, M., Steel, D. & Browne, W.J., 2014. Multiple Membership Multiple Classification Models for Social Network and Group Dependences. *Journal of the Royal Statistical Society (Series A) (Statistics in Society)*, 177(2), pp.439–455.

TripAdvisor, 2018. London Restaurants. Available at: http://www.tripadvisor.co.uk/Restaurants-g186338-London_England.html [Accessed May 16, 2018].

Trochim, W.M.K., Donnelly, J.P. & Arora, K., 2016. *Research Methods: The Essential Knowledge Base* 2nd Editio., Cengage Learning.

Trowler, V., 2010. Student engagement literature review. *Higher Education*, (November), pp.1–15.

Tsao, W.-C. et al., 2015. Compliance with eWOM: The influence of hotel reviews on booking intention from the perspective of consumer conformity. *International Journal of Hospitality Management*, 46, pp.99–111.

Turner, J.C., 1984. Social identification and psychological group formation. In H. social dimension-E. developments in social psychology vol 2/25 social_identification_and_psychological_group_formation. pd. Tajfel, ed. *The Social Dimension: European Developments in Social Psychology*. European Studies in Social Psychology. Cambridge: Cambridge University Press, pp. 518–538.

Turner, J.C. & Reynolds, K.J., 2012. Self-categorization theory. In *Handbook of Theories of Social Psychology*. pp. 399–417.

Turner, V. et al., 2014. The Digital Universe of Opportunities: Rich Data and Increasing Value of the Internet of Things. *IDC White Paper*, (April). Available at: http://www.emc.com/leadership/digital-universe/2014iview/it-imperatives.htm [Accessed April 13, 2015].

Tversky, A. & Kahneman, D., 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science*, 185(4157), pp.1124–1131.

Tversky, A. & Kahneman, D., 1981. The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), pp.453–458.

Vahl, A., 2011. 4 Things You Must Know About Facebook Smart Lists. Available at: https://www.andreavahl.com/facebook/4-things-you-must-know-about-facebook-smart-lists.php [Accessed April 18, 2018].

Verkhivker, A., 2018. Zuckerberg Testifies Before Congress And Now Everyone Is Worried Tech’s Too Dominant. *Forbes*.

Wall, T.D. et al., 1986. Outcomes of Autonomous Workgroups : A Long Term Field Experiment. *Academy of Management Journal*, 29(2), p.pp.280-304.

Wasserman, S. & Faust, K., 1994. *Social Network Analysis: Methods and Applications* illustrate., Cambridge University Press.

Weick, K.E., 1969. *The social psychology of organizing*, Reading, Massachusetts: Addison-Wesley.

Whetten, D.A., 1989. What Constitutes a Theoretical Contribution? *Academy of Management Review*, 14(4), pp.490–495.

Whittaker, E. & Kowalski, R.M., 2015. Cyberbullying Via Social Media. *Journal of School Violence*, 14(1), pp.11–29.

Wilkinson, D. & Thelwall, M., 2012. Trending Twitter Topics in English: An International Comparison. *Journal of the American Society for Information Science and Technology*, 63(8), pp.1631–1646.

de Winter, J.C.F. & Dodou, D., 2010. Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, 15(11), pp.1–16.

Wong, J.C., 2018. I was one of the first people on Facebook. I shouldn't have trusted Mark Zuckerberg. *The Guardian*.

Worldometers, 2018. Current World Population. Available at: http://www.worldometers.info/world-population/ [Accessed January 10, 2017].

Wu, S. et al., 2011. Who Says What to Whom on Twitter. *Proceedings of the 20th International Conference on World Wide Web*, pp.705–714.

YouTube, 2009. Five Stars Dominate Ratings. *YouTube Official Blog*. Available at: https://youtube.googleblog.com/2009/09/five-stars-dominate-ratings.html [Accessed March 13, 2017].

Yu, C.H., 2005. Misconceived relationships between logical positivism and quantitative research. *Philosophical Foundations of Quantitative Research Methodology*, 1(1), pp.p1-39.

Zhao, S., Grasmuck, S. & Martin, J., 2008. Identity construction on Facebook: Digital empowerment in anchored relationships. *Computers in Human Behavior*, 24(5), pp.1816–1836.

Zimbardo, P.G., 1969. The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation*, pp.237–307.

APPENDICES

APPENDIX 3.1 – QUASI-EXPERIMENTAL DESIGN IN PEERWISE

This appendix presents all the documented changes that were done in PeerWise for the selected module of The University of Sheffield.

Quasi-experimental set-up

As outlined in the Methodology, the quasi-experiments were designed as follows:

Group 1

Autumn semester 2014-15

[Baseline]

Group 2

Autumn semester 2015-16

[Make all users identifiable]

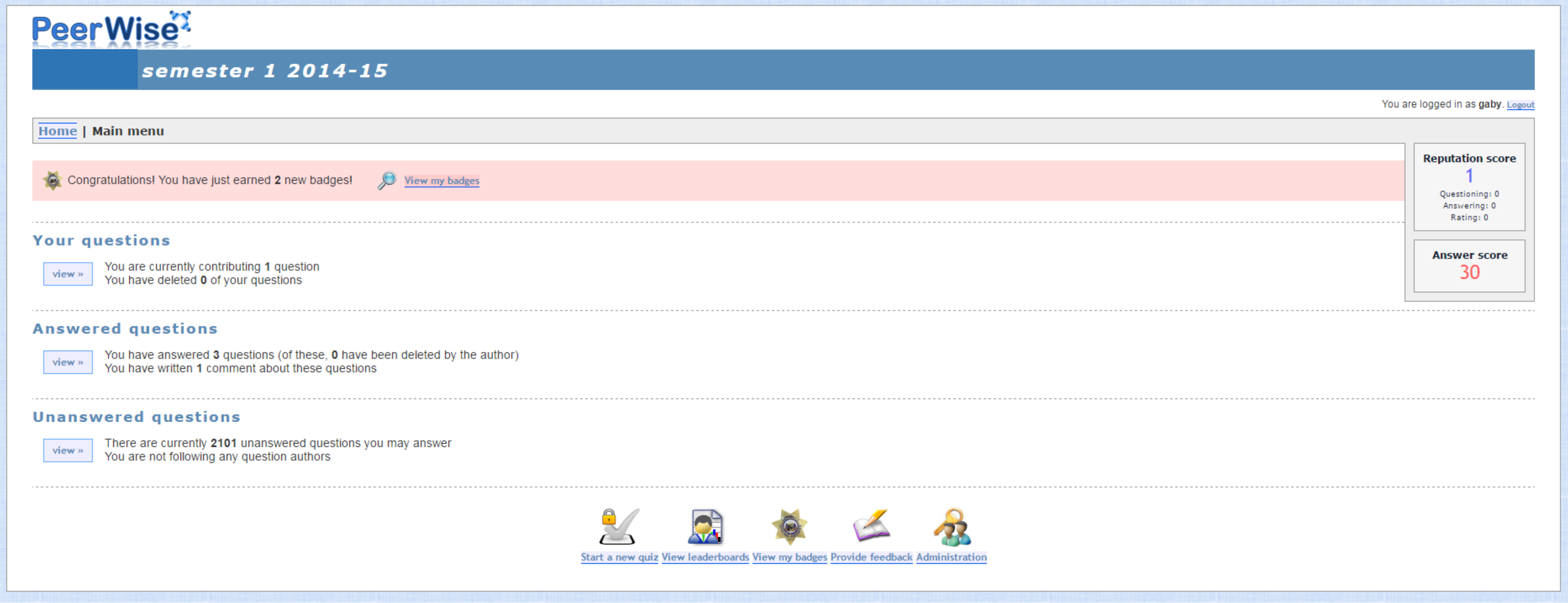
Group 3

Autumn semester 2016-17

[Change in rating scale]

Group 1 – Baseline: anonymous users & likert scale

The baseline year is basically how PeerWise is set up for all students worldwide. The following image shows the main menu of the site. When a student has just joined PeerWise, she has two main options: to author a question (in "Your questions") or to answer an available one (in "Unanswered questions"). Once the student has authored one or more questions, she can follow them up in "Your questions". Likewise, once she has answered one or more questions, she can refer back to them through "Answered questions".



When a student sets up a new question, she writes the question's text and then the software asks how many multiple-choice answers she wants to give. She can choose any number between two and five. After this, she must indicate which is the 'correct' answer. Finally, she has a space for Explanation" in case she wants to give some clarification to her peers on why the determined answer is correct.

Milton Friedman argues that the corporate executive should avoid:

Alternatives

OPTION	ALTERNATIVE
A	Spending someone else's money for a social interest
B	Spending his or her own money for a social interest
C	Spending his or her own money for a social interest
D	Using money earned from business to fund social causes
E	Having opinions about social justice issues

Explanation

The following explanation has been provided relating to this question:

In the textbook on P.85 it writes that "... the corporate executive would be spending someone else's money for a general social interest.Insofar as his actions in accord with his social responsibility reduce returns to stockholders, he is spending their money. Insofar as his actions raise the price to customers, he is spending the customers money. Insofar as his actions lower the wages of some employees, he is spending their money.
For Friedman this process raises political questions on two levels: principle and consequences. On the level of political principle, the imposition of taxes and the expenditure of tax proceeds are governmental functions. Here the corporate executive, as an agent of shareholders, acts as if he were simultaneously the legislator, executive, and jurist. He is to decide whom to tax by how much and for what purpose, and he is to spend the proceeds; all this guided only by general exhortations to restrain price inflation, improve the environment, fight poverty, and so on, and on.
The whole justification for permitting the corporate executive to be selected by the shareholders is that the executive is an agent serving the interests of his principal. This justification disappears when the corporate executive imposes taxes and spends the proceeds for social purposes. He becomes, in effect, a public employee, a civil servant; even though he remains in name an employee of a private enterprise.
On grounds of political principle, it is intolerable that such civil servants V insofar as their actions in the name of social responsibility are real and not just window-dressing V should be selected as they are now. If they are to be civil servants, then they must be selected or elected through a political process. If they are to impose taxes and make expenditures to foster social objectives, then political machinery must be set up to make the assessment of taxes and to determine through a political process the objectives to be served. This forms the basis for Friedman's argument that CSR is wrong. First, it involves the acceptance of relying on political mechanisms, not market mechanisms, as the appropriate way to determine the allocation of scarce resources to alternative uses. Second, the corporate executive is not a properly appointed public employee and therefore not politically accountable.
On the grounds of consequences, it is not evident that the corporate executive can properly discharge his alleged social responsibilities He is presumably an expert in running his company. But nothing about his selection makes him an expert on inflation, improving the environment, fighting poverty, and so on. Even if he happens to be an expert on these subjects, how much cost is he justified in imposing on his shareholders, customers and employees for this social purpose? What is his appropriate share and what is the appropriate share of others? These are political questions that he is not empowered to decide upon for the public interest.
And, whether he wants to or not, can he get away with spending his shareholders, customers or employees money? Might not the shareholders fire him? His customers and his employees can desert him for other producers and employers less scrupulous in exercising their social responsibilities."

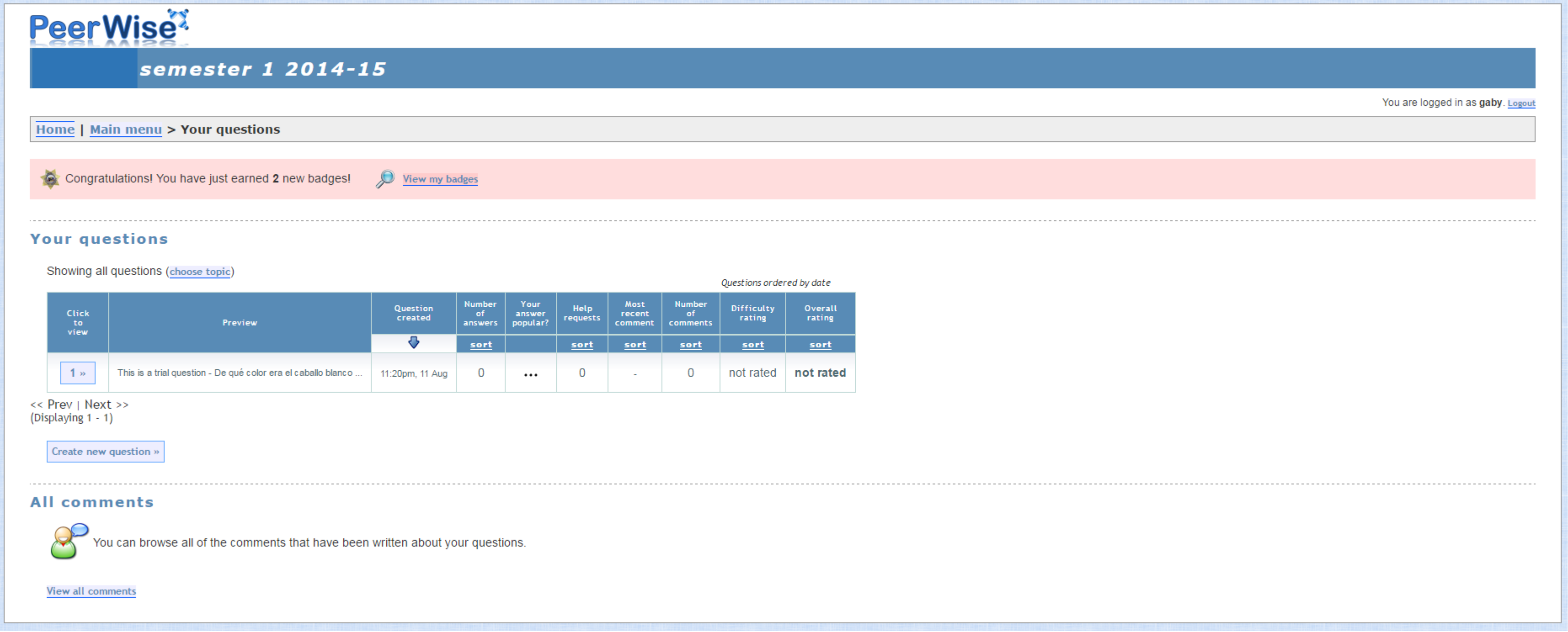
Request help
Improve explanation

Topics

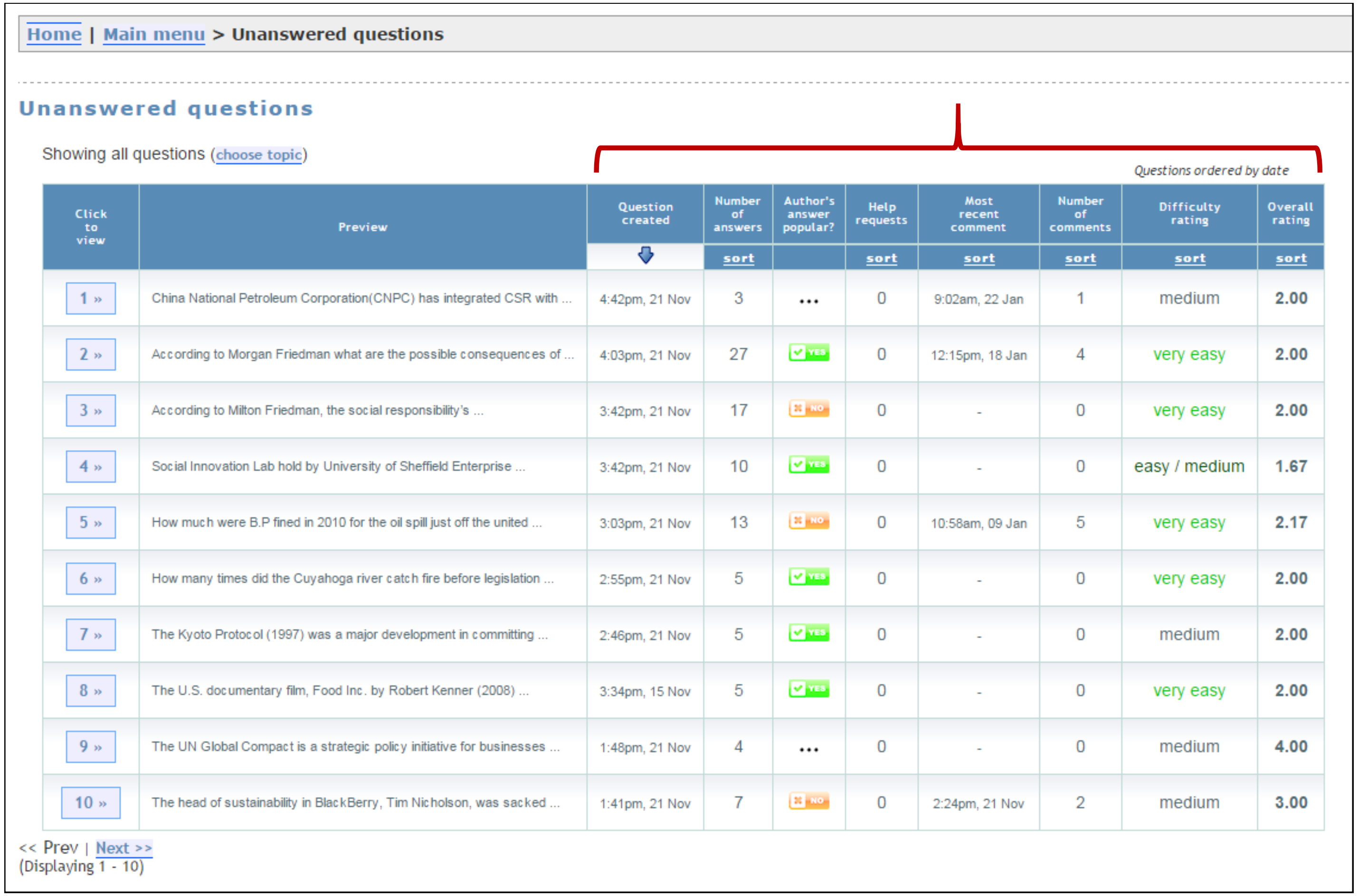
The following topics have been indicated as being relevant to this question:

Milton Friedman

Later, in the "Your questions" section, students can manage (edit or delete) the questions they authored, or create new ones. Moreover, they can check the average rating score for each question, and how many people have targeted them. Finally, they can also check if someone has posted a comment on any of their questions.



Furthermore, if students want to target their peers' questions, they go to "Unanswered questions". In here, questions were usually arranged by (latest) created question, but users could also sort and choose to answer questions based on: number of answers, popular answer, help requests, most recent comment, number of comments, difficulty rating and overall rating.



Home | Main menu > Unanswered questions

Unanswered questions

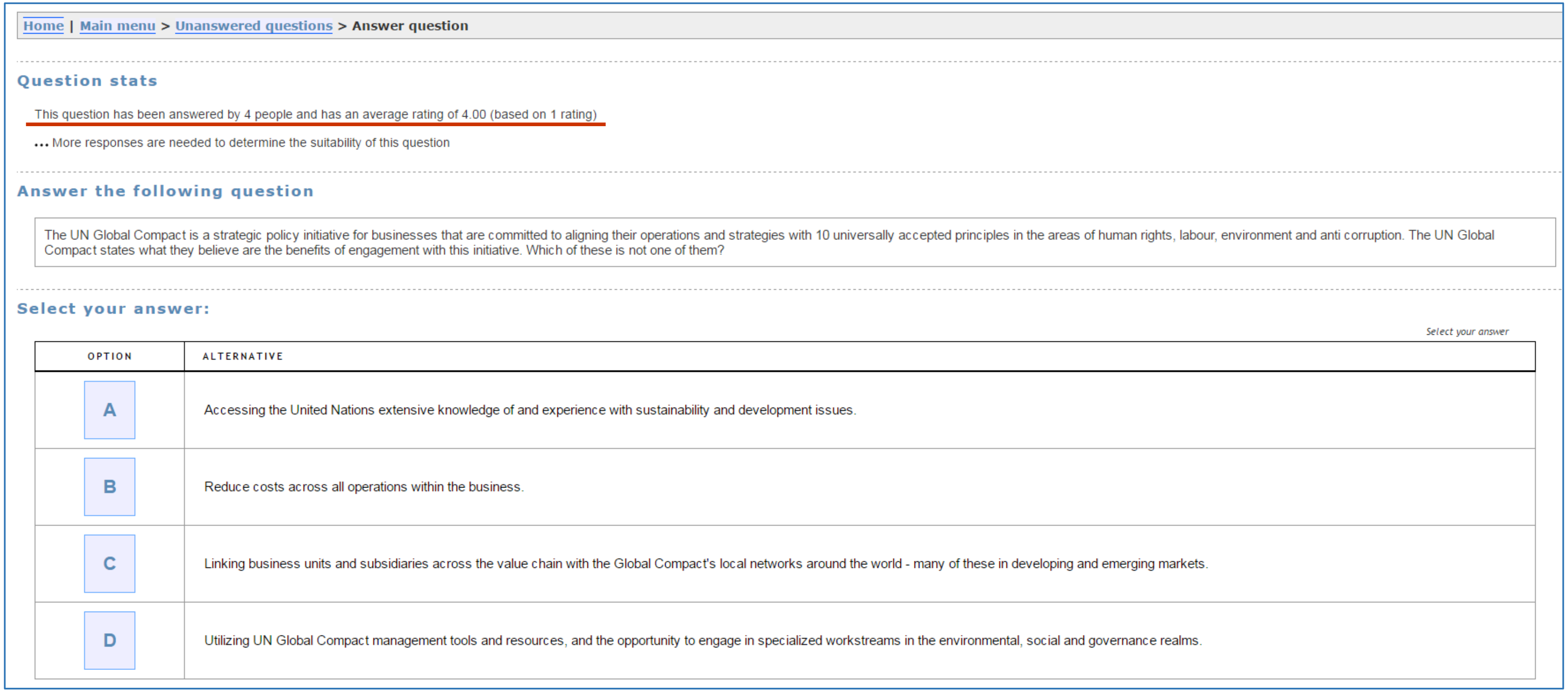
Showing all questions (choose topic)

Questions ordered by date

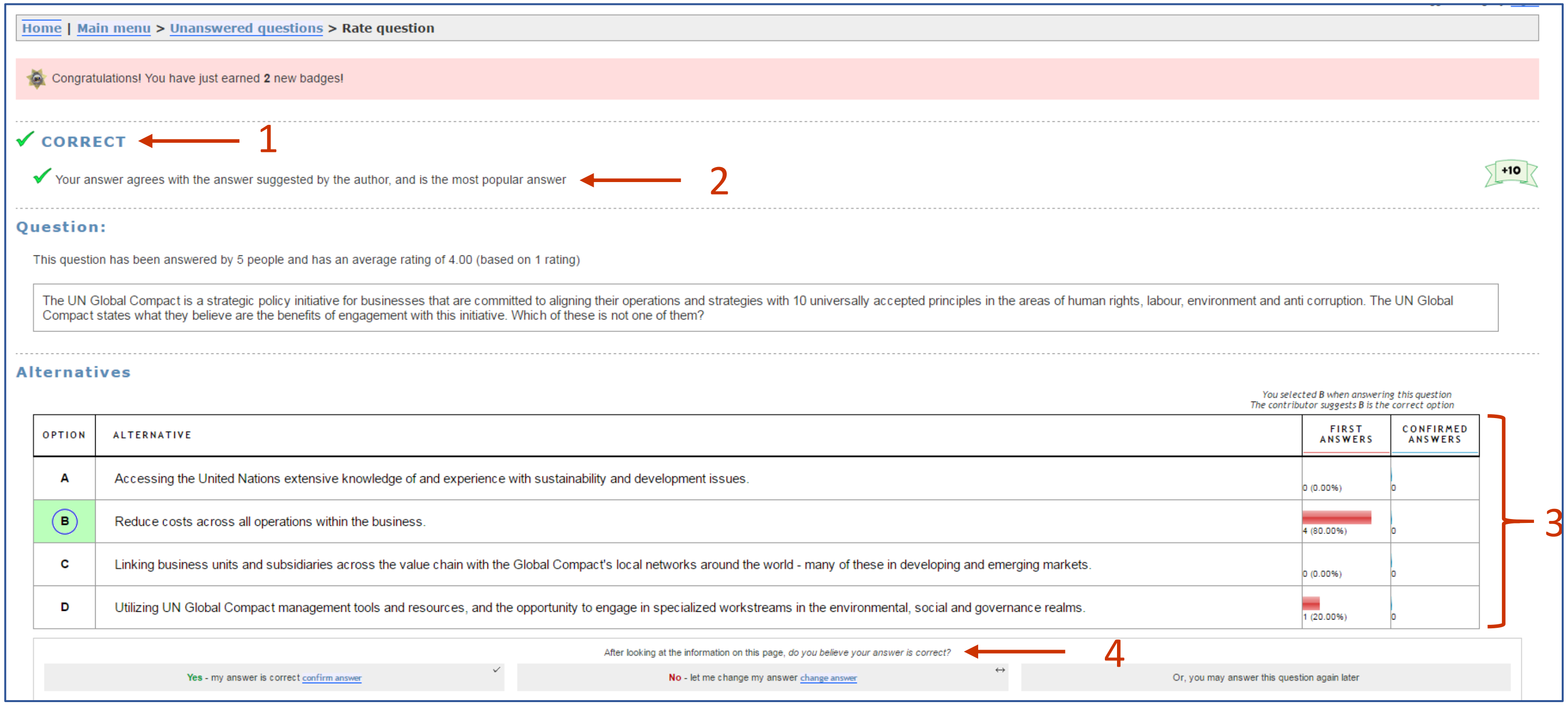
Click to view	Preview	Question created	Number of answers	Author's answer popular?	Help requests	Most recent comment	Number of comments	Difficulty rating	Overall rating
			sort		sort	sort	sort	sort	sort
1 »	China National Petroleum Corporation(CNPC) has integrated CSR with ...	4:42pm, 21 Nov	3	...	0	9:02am, 22 Jan	1	medium	2.00
2 »	According to Morgan Friedman what are the possible consequences of ...	4:03pm, 21 Nov	27	YES	0	12:15pm, 18 Jan	4	very easy	2.00
3 »	According to Milton Friedman, the social responsibility's ...	3:42pm, 21 Nov	17	NO	0	-	0	very easy	2.00
4 »	Social Innovation Lab hold by University of Sheffield Enterprise ...	3:42pm, 21 Nov	10	YES	0	-	0	easy / medium	1.67
5 »	How much were B.P fined in 2010 for the oil spill just off the united ...	3:03pm, 21 Nov	13	NO	0	10:58am, 09 Jan	5	very easy	2.17
6 »	How many times did the Cuyahoga river catch fire before legislation ...	2:55pm, 21 Nov	5	YES	0	-	0	very easy	2.00
7 »	The Kyoto Protocol (1997) was a major development in committing ...	2:46pm, 21 Nov	5	YES	0	-	0	medium	2.00
8 »	The U.S. documentary film, Food Inc. by Robert Kenner (2008) ...	3:34pm, 15 Nov	5	YES	0	-	0	very easy	2.00
9 »	The UN Global Compact is a strategic policy initiative for businesses ...	1:48pm, 21 Nov	4	...	0	-	0	medium	4.00
10 »	The head of sustainability in BlackBerry, Tim Nicholson, was sacked ...	1:41pm, 21 Nov	7	NO	0	2:24pm, 21 Nov	2	medium	3.00

<< Prev | Next >>
(Displaying 1 - 10)

In order to expand on some of the above mentioned options for sorting questions, the author will explain what happens when answering a question. Firstly, once a question is selected, the student is presented with the question and the possible answers. A short summary at the top indicates how many people have answered and rated the question, and if applicable, the average rating is shown. However, it should also be noted that both the author of the question and the students that have previously targeted and rated the question, are all kept anonymous in PeerWise's original design:

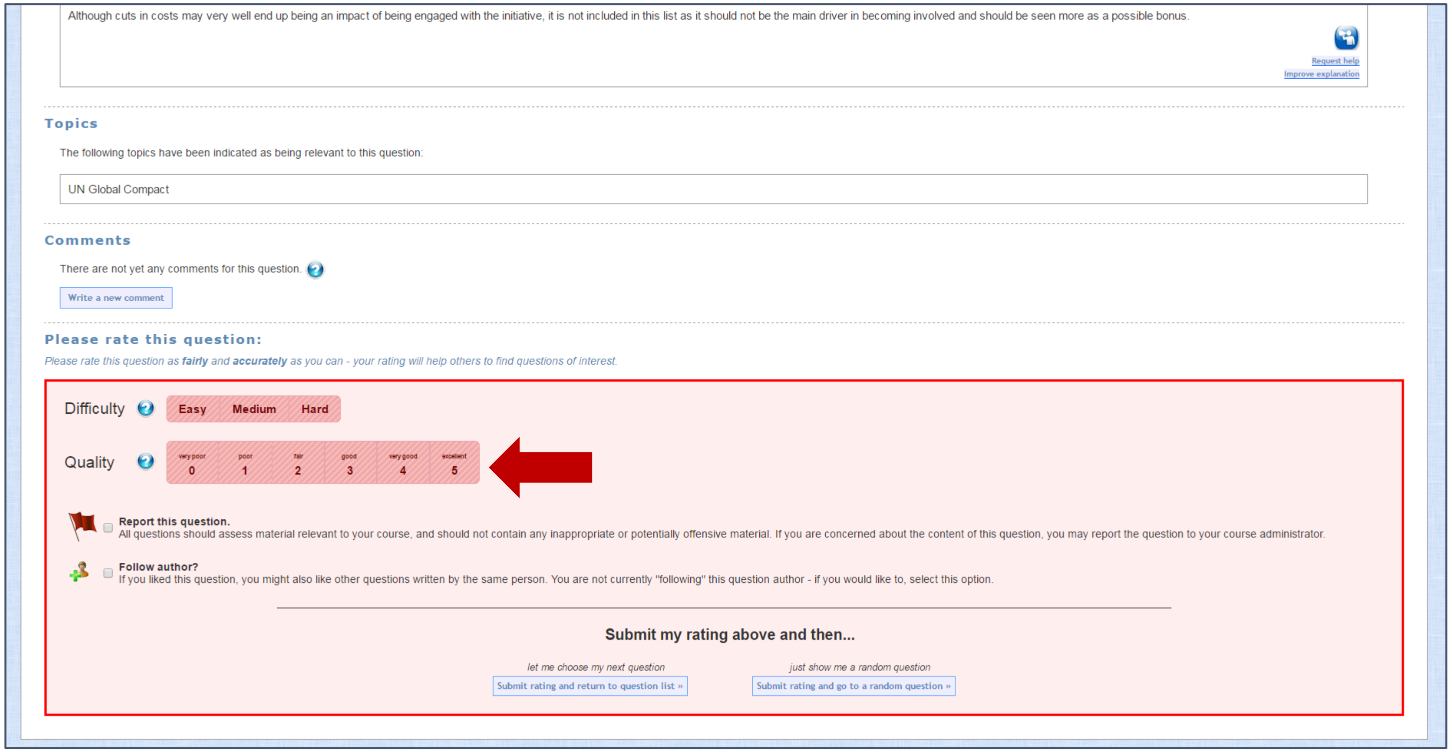


Afterwards, the student must select an answer from the available ones. Immediately after the answer is selected, the software gives feedback to the student. This shows: 1) If the chosen answer is correct or incorrect; 2) If the chosen answer is the most 'popular' one (i.e., the most selected by other students who have targeted the same question); 3) The percentage of peers who have answered and confirmed all of the available answers; and finally 4) It gives the student the option of changing or confirming her answer. Please note that it is possible to change the answer and/or tackle questions more than once. However, <u>this research will only analyse the very first answer</u> that students gave. Also it is worth reiterating students' grades did not take into account the amount of right or wrong questions; on the contrary, students were encouraged to be critical and to disagree if they thought a specified 'correct' answer was in fact wrong.



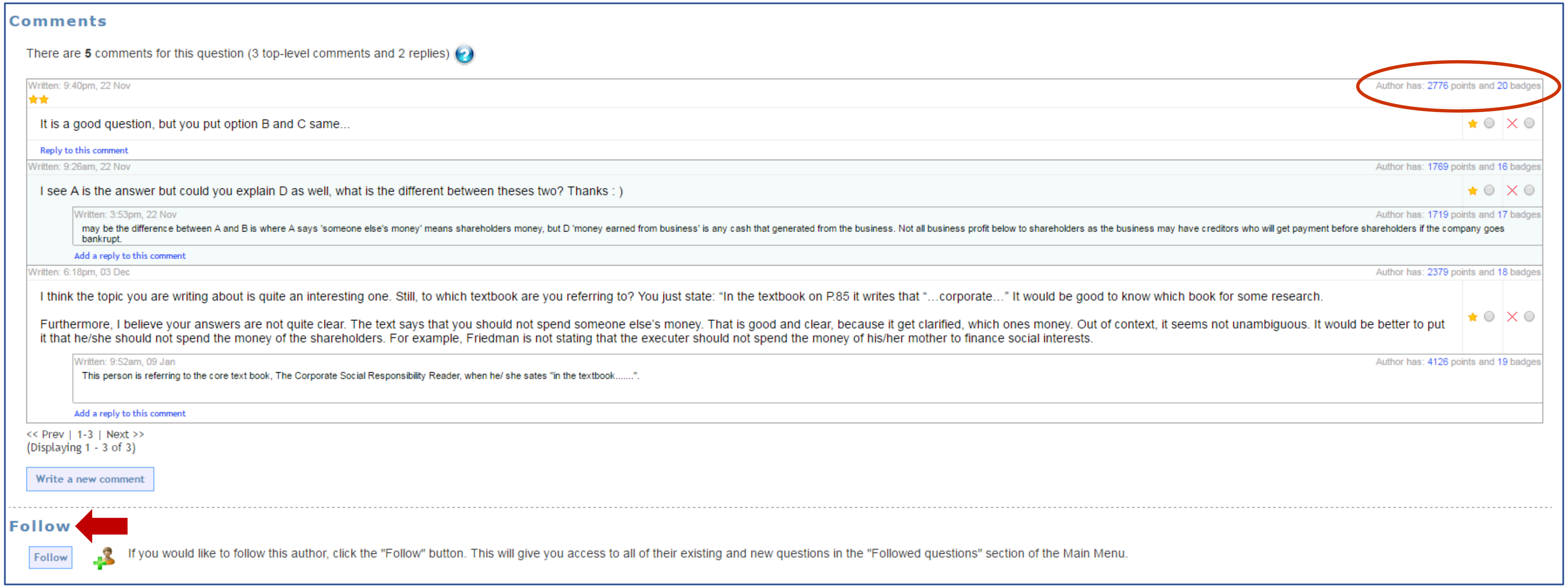
After the answer is confirmed, the author's explanation of the answer appears (if applicable, as this was optional for authors). At this point, the student who just targeted the question is given the options of leaving a comment and/or rating the question (in terms of difficulty and quality). The 'quality' rating is the one that determines the question's rating average, which is used to sort questions and is a component for getting badges and a criteria for being in the leaderboards.

Commenting and rating are optional, both for PeerWise and for the module's assessment.

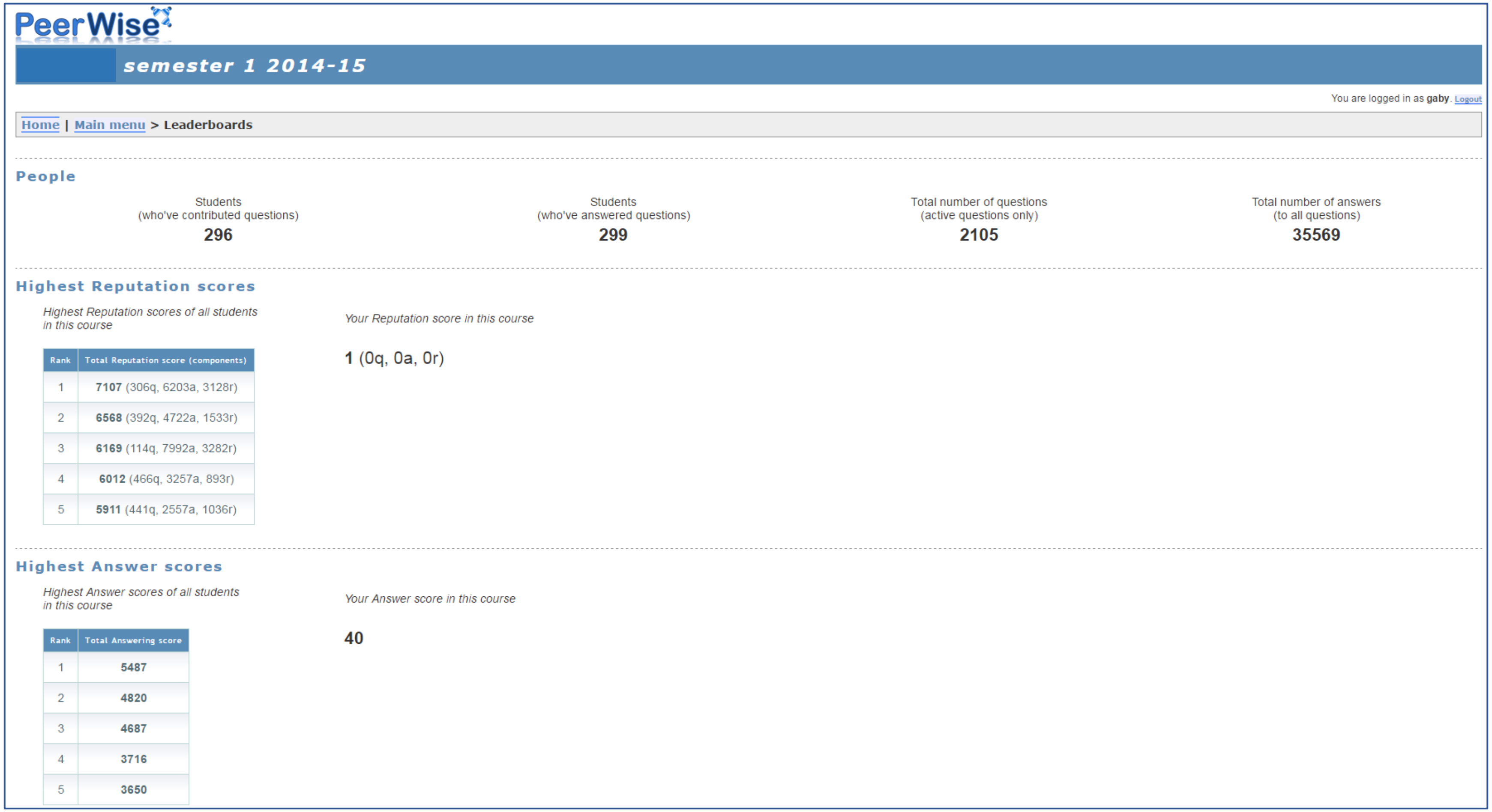


Also, once the student has submitted an answer (and regardless of whether she commented or rated the question), she is now able to see all the comments that other users have left. Moreover, once a student has targeted a question this will appear in “Answered questions” and they can go back and leave comments at any point. The same applies for “following” an author, which can be done at any time.

Note that PeerWise was initially set-up so that both commenting and rating are anonymous.



Finally, PeerWise has in place leaderboards that all students can access. These were also anonymous and allowed students to compare themselves with those with the top-highest scores in a number of categories:



Top rated questions

Top 5 rated questions for this course (rated by at least 5 students)

Rank	Question rating
1	4.0800
2	4.0000
3	4.0000
4	4.0000
5	4.0000

Highest rating of any of your questions (rated by at least 5 students)

-

Most questions answered

Most questions answered by a single student

Rank	Questions answered
1	683
2	615
3	601
4	528
5	527

Number of questions you have answered

4

Most "answered" question contributor

Total number of answers to all questions contributed by a single student

Rank	Total number of answers
1	694
2	691
3	596
4	577
5	501

Total number of answers to all questions you have contributed

0

Most "agreed with" critic

Sum of agreement ratings of all comments written by a single student

Rank	Agreement with comments
1	235
2	210
3	190
4	91
5	71

Sum of agreement ratings of all comments written by you

0

Most "followed" question author

Number of followers of a single author

Rank	Total number of followers
1	22
2	20
3	19
4	19
5	18

Number of students following your questions

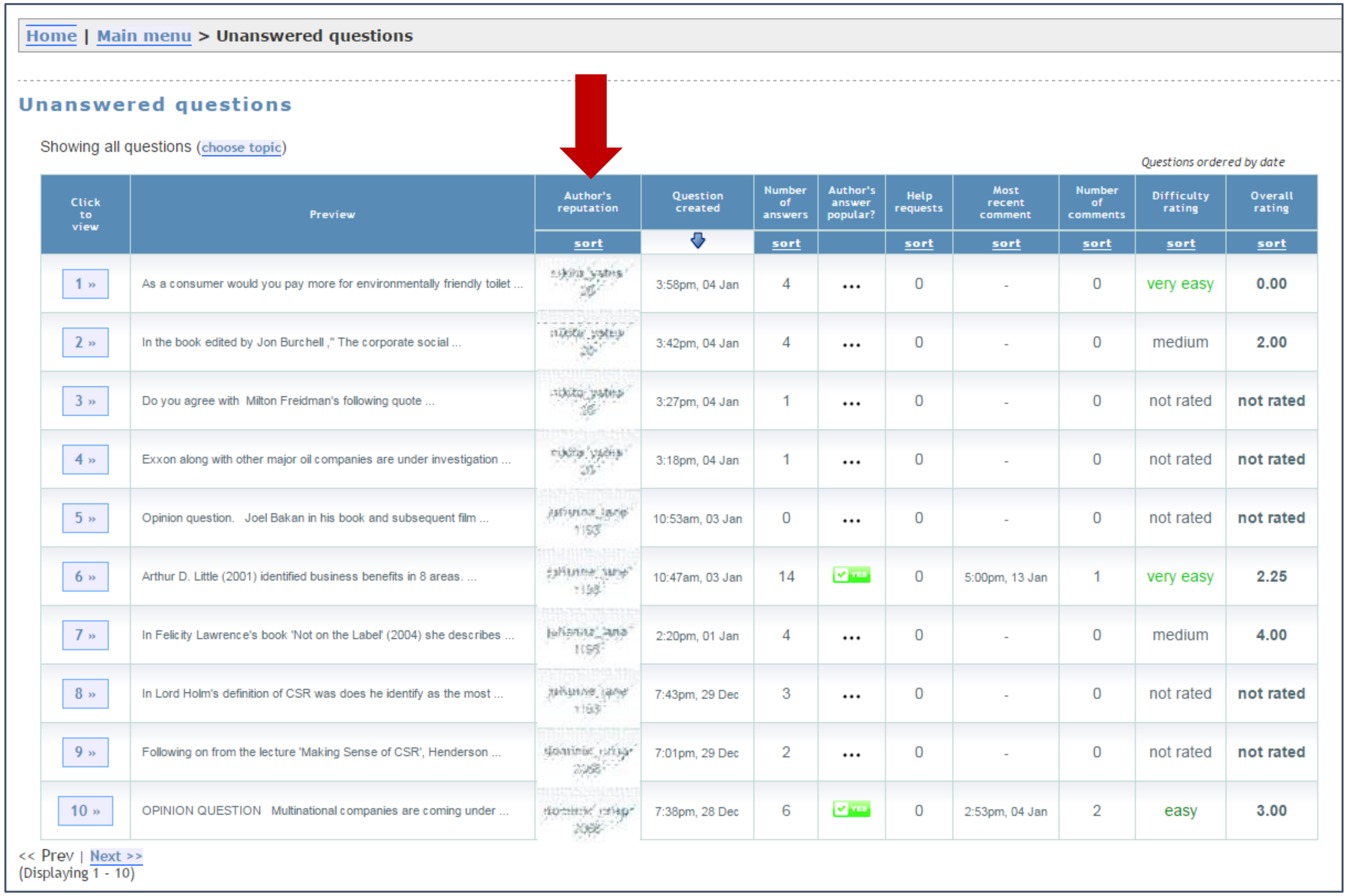
0

Your usage

- Number of days you have actively participated: **3**

Group 2 – Removing anonymity: identifiable users & likert scale

The first change of being identifiable was that questions could now be sorted by 'Author's reputation' (which included the author's name).



Home | Main menu > Unanswered questions

Unanswered questions

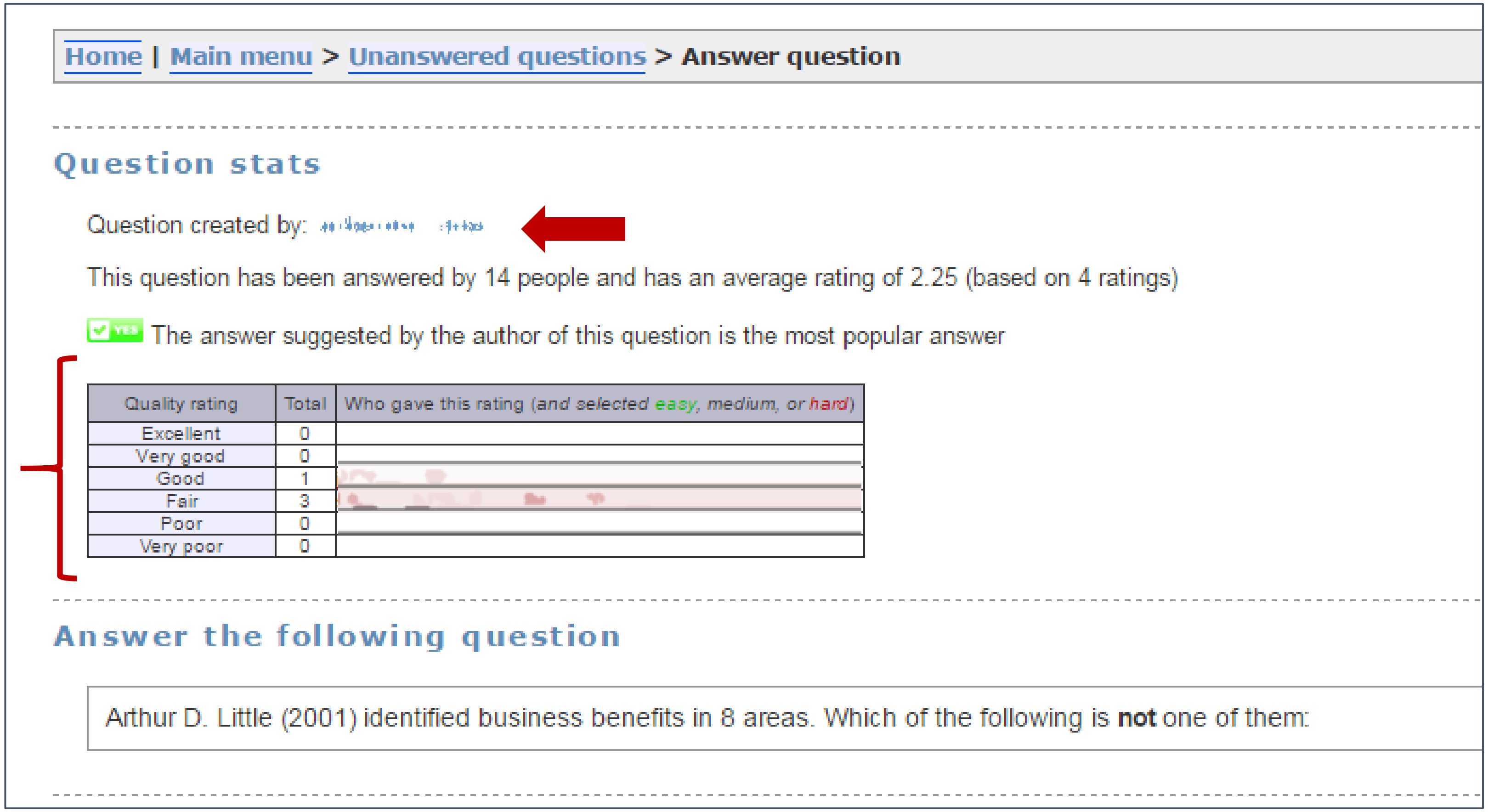
Showing all questions (choose topic)

Questions ordered by date

Click to view	Preview	Author's reputation	Question created	Number of answers	Author's answer popular?	Help requests	Most recent comment	Number of comments	Difficulty rating	Overall rating
		sort		sort		sort	sort	sort	sort	sort
1 »	As a consumer would you pay more for environmentally friendly toilet ...	[illegible]	3:58pm, 04 Jan	4	...	0	-	0	very easy	0.00
2 »	In the book edited by Jon Burchell ," The corporate social ...	[illegible]	3:42pm, 04 Jan	4	...	0	-	0	medium	2.00
3 »	Do you agree with Milton Freidman's following quote ...	[illegible]	3:27pm, 04 Jan	1	...	0	-	0	not rated	not rated
4 »	Exxon along with other major oil companies are under investigation ...	[illegible]	3:18pm, 04 Jan	1	...	0	-	0	not rated	not rated
5 »	Opinion question. Joel Bakan in his book and subsequent film ...	[illegible]	10:53am, 03 Jan	0	...	0	-	0	not rated	not rated
6 »	Arthur D. Little (2001) identified business benefits in 8 areas. ...	[illegible]	10:47am, 03 Jan	14	YES	0	5:00pm, 13 Jan	1	very easy	2.25
7 »	In Felicity Lawrence's book 'Not on the Label' (2004) she describes ...	[illegible]	2:20pm, 01 Jan	4	...	0	-	0	medium	4.00
8 »	In Lord Holm's definition of CSR was does he identify as the most ...	[illegible]	7:43pm, 29 Dec	3	...	0	-	0	not rated	not rated
9 »	Following on from the lecture 'Making Sense of CSR', Henderson ...	[illegible]	7:01pm, 29 Dec	2	...	0	-	0	not rated	not rated
10 »	OPINION QUESTION Multinational companies are coming under ...	[illegible]	7:38pm, 28 Dec	6	YES	0	2:53pm, 04 Jan	2	easy	3.00

<< Prev | Next >>
(Displaying 1 - 10)

Moreover, when rating a question the changes were as follows: 1) The name of the author (i.e. Name_LastName) would be presented. As well as a breakdown of all given ratings, also showing the names of students.


Home | Main menu > Unanswered questions > Answer question

Question stats

Question created by:

This question has been answered by 14 people and has an average rating of 2.25 (based on 4 ratings)

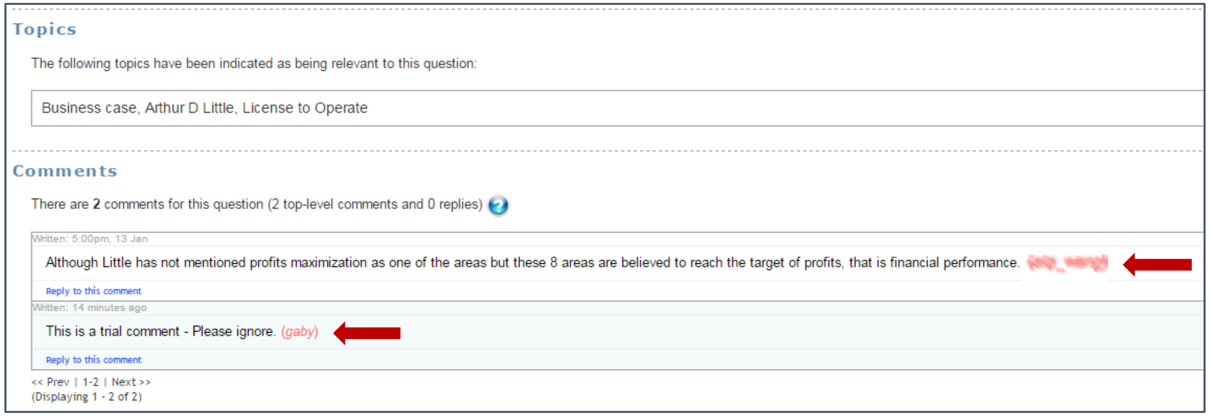
The answer suggested by the author of this question is the most popular answer

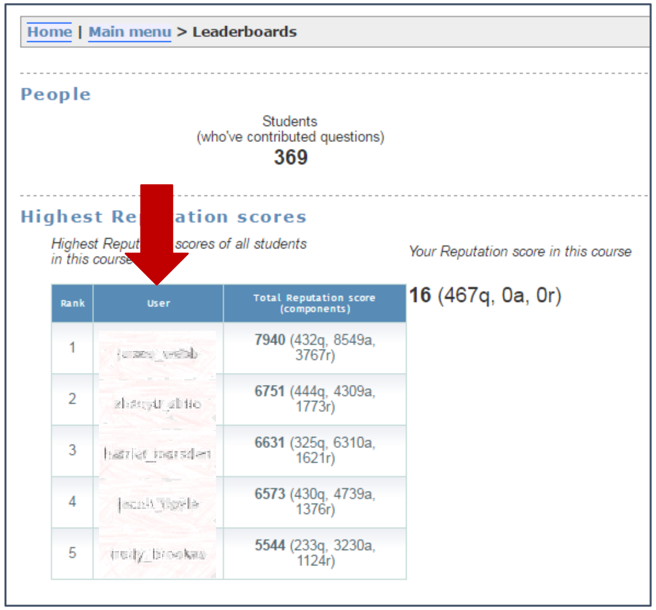
Quality rating	Total	Who gave this rating (*and selected easy, medium, or hard*)
Excellent	0	
Very good	0	
Good	1	
Fair	3	
Poor	0	
Very poor	0	

Answer the following question

Arthur D. Little (2001) identified business benefits in 8 areas. Which of the following is **not** one of them:

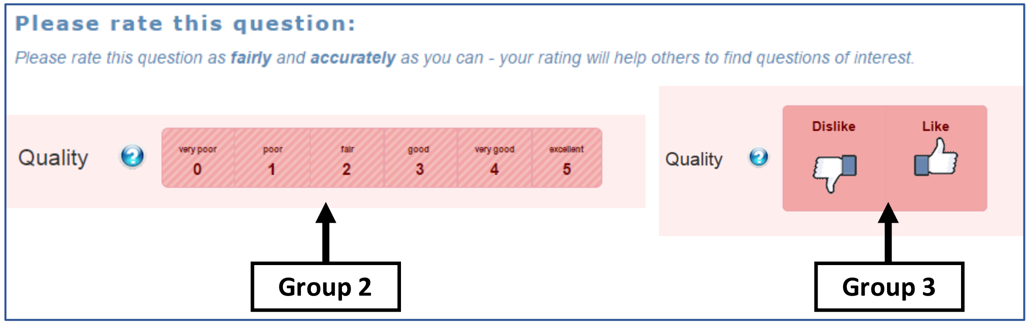
Further, commenting and leaderboards also became identifiable:

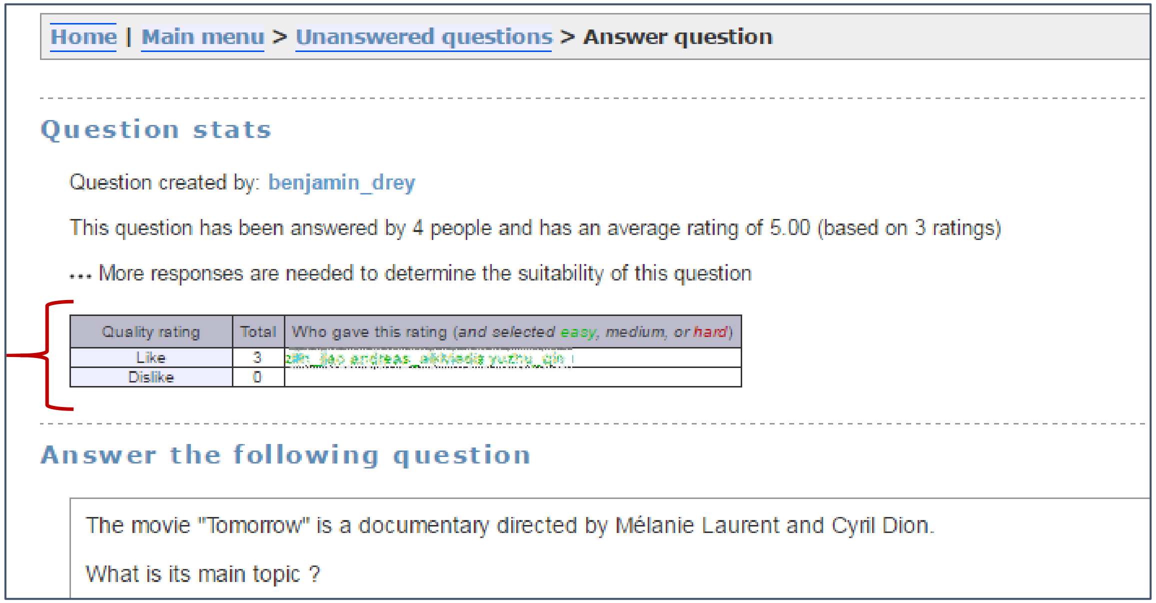




Group 3 – Changing the rating scale: identifiable users & dichotomous scale

The third group experienced a change in rating scale. Other than the dichotomous scale (which affected the rating breakdown) the rest of the quasi-experimental design was identical to that of Group 2.





APPENDIX 3.2 – QUESTIONNAIRE (GROUP 3)

Thank you for answering this questionnaire! It should take you 10 minutes to answer and you will automatically be considered for a raffle of one £50 and five £20 Amazon vouchers! (six Amazon vouchers, worth £150 in total, only for your MGT357 module).

This questionnaire is part of the the *"Understanding the Flow of Information within User Generated Content in Virtual Education Learning Environments"* research. that was introduced to you in the first lecture. It is NOT part of your module and therefore: is optional, will NOT impact your grade in any way, and your answers will NOT be shared with the teaching staff on the module. Moreover, you will be assigned an anonymous code which will be used throughout the analysis process, and therefore all results will be reported anonymously. This project has been reviewed and approved by the *University of Sheffield Management School Ethics Committee* in accordance with the University of Sheffield ethics policy.

By answering this questionnaire, I agree to take part in the study and I understand that my answers will be used for research purposes in an anonymous way.

☐ Yes

Rating Scale

Overall, how helpful do you think PeerWise's rating scale is? (*PeerWise's rating scale currently allows you to choose from two options when rating the quality of a question: Like and Dislike*). Do you consider these two options adequate for assessing the quality of questions?

	Very Helpful	Helpful	Neutral	Unhelpful	Very Unhelpful
PeerWise's rating scale	○	○	○	○	○

If you could vote regarding whether to keep the rating scale or change it for another one, would you rather:

- Keep it as it is
- Change it - give more options (i.e. have a rating scale from 1 to 5)
- Change it - give fewer options (i.e. only ‘Like’)

Please explain why, if you wish (Optional)

Which of the following factors did **<u>you</u>** take into account when rating a question? (Select all that apply)

- Whether you knew the author personally (i.e. if you were friends with the person who authored the
- question)
- The overall rating that the question already had (i.e. how others had rated the question before you)
- Your perceived quality of the question (i.e. if the content of the question)
- The author's reputation score in PeerWise
- Whether the author of the question was featured in the PeerWise leaderboards
- The popularity of the question (i.e. the number of answers or comments that the question had)

Which of the following factors do you think **<u>other students</u>** took into account when rating a question? (Select all that apply)

- The author's reputation score in PeerWise
- Their perceived quality of the question (i.e. if the content of the question)
- The popularity of the question (i.e. the number of answers or comments that the question already had)
- Whether they knew the author personally (i.e. if they were friends with the person who authored the question)
- Whether the author of the question was featured in PeerWise leaderboards
- The overall rating that the question already had (i.e. how others had rated the question before them)

Overall, would you say that the five questions you authored were rated:

	Mostly Fairly	Sometimes Fairly	Neutral	Sometimes Unfairly	Mostly Unfairly
My 5 questions were rated	o	o	o	o	o

Please explain further, if you wish (Optional)

Use of Identifiers

How helpful was to identify your classmates in PeerWise by their usernames (name_lastname)?

	Very Helpful	Helpful	Neutral	Unhelpful	Very Unhelpful
Use of identifiable usernames	o	o	o	o	o

If you could have a say regarding your classmates' identities within PeerWise, would you rather:

- Have them identifiable (i.e. be able to identify my classmates' questions and actions)
- Have them anonymous (i.e. I would not like to know my classmates' identities)

If you could choose, would you rather:

- Be identifiable (i.e. keep using 'name_lastname' as my username, so others can identify my questions and actions)
- Be anonymous (i.e. use any other username that doesn't allow others to identify me)

Please expand your answer regarding being identifiable VS anonymous, if you wish (Optional)

PeerWise associates a number of elements to your username (i.e. author's reputation, badges and leaderboards).To what extent did these motivate you to author and answer questions?

	Highly Motivated	Motivated	Neutral	Demotivated	Highly Demotivated
Reputation, badges & leaderboards	○	○	○	○	○

Use of PeerWise

PeerWise allowed you to sort questions in different ways in order for you to select which ones you wanted to answer. From the list below, please select which ones you used most frequently for sorting and selecting questions (mark <u>all</u> that apply)

- ☐ I would select questions based on the topic
- ☐ I was following some authors in PeerWise, and would select their questions first
- ☐ I would sort and select questions by popularity (number of answers or number of comments)
- ☐ I would sort and select questions based on the author's reputation
- ☐ I would sort and select questions based on their difficulty
- ☐ I would sort and select questions based on their overall rating
- ☐ I would randomly select questions
- ☐ I would select questions from the people I knew in the class

Which was your main method for selecting questions? (please select <u>only one</u>)

- ○ I would sort and select questions based on their overall rating
- ○ I would sort and select questions based on their difficulty
- ○ I would sort and select questions by popularity (number of answers or number of comments)
- ○ I would select questions from the people I knew in the class
- ○ I would select questions based on the topic
- ○ I would randomly select questions
- ○ I would sort and select questions based on the author's reputation
- ○ I was following some authors in PeerWise, and would select their questions first

Did you use the “Follow" (author) function?

- Yes
- No

(Follow-up, if yes) Would you say:

- I mostly followed people I personally knew
- I mostly followed people based on the quality of their questions

Please name at least three (but as many as you like) people from this class that you know. For multiple selection: press 'Ctrl' in PC, 'cmd' in Mac, or just scroll and select in a mobile phone. *(This information will help us understand from whom students learnt the most)*

Name_LastName1
Name_LastName2
Name_LastName3
Name_LastName4
Name_LastName…
Name_LastName309

Could you please tell us how frequently you interact with the above-mentioned classmates, outside of the class? (*This information would help us to further understand from whom students learnt the most, but if you prefer not to answer, simply choose “N/A”*).

Frequently: *Classmates that you regularly see outside your class, at least once a week (and if you use social media, they are within your social network).*
Occasionally: *Classmates that you sporadically see outside your class, but you have met them outside the classroom at least once in the last year and/or you know updates from their lives through social media.*
Only in class: *Classmates you only know from your class, but you have never seen them outside the class and you don’t have them in social media.*

	Frequently	Occasionally	Only in class	N/A
Name_LastName1	○	○	○	○
Name_LastName2	○	○	○	○
Name_LastName3	○	○	○	○

Learning with PeerWise

Overall, how helpful do you consider the use of PeerWise in a learning environment (i.e. that authoring questions and engaging in debates helped you learn)?

	Very Helpful	Helpful	Neutral	Unhelpful	Very Unhelpful
Learning with PeerWise	○	○	○	○	○

Please expand your answer, if you wish (Optional) - *If you find anything valuable about authoring questions and engaging in debates, we would like to know it. Also if this wasn't the case.*

Demographic Questions (all optional)

Gender (Optional)

- Male
- Female
- Other

Age (Optional)

Nationality (Optional)

Which of the following social networking sites do you use? Mark all that apply (Optional)

- ☐ Facebook
- ☐ Twitter
- ☐ Snapchat
- ☐ Pinterest/Instagram
- ☐ Weibo
- ☐ QZone
- ☐ Others

If others, please mention which ones (Optional)

How long would you say you spend every day in social networking sites? (Optional)

- ○ I don't use social networking sites
- ○ I get notifications throughout the day, so it varies a lot
- ○ Between 5 and 30 minutes
- ○ Between 30 minutes and an hour
- ○ Between 1 and 2 hours
- ○ Between 2 and 3 hours
- ○ More than 3 hours

Focus groups

Would you be willing to attend a later focus group (group discussion on Peerwise)?

- ○ Yes - Email me
- ○ Maybe, depending on the dates and the incentives - Email me
- ○ No - Don't email me

Thank you very much for your time and feedback!!

The results of the raffle will be posted on MOLE at the end of the year - And if you are a winner of one of the vouchers, you will also receive an email with the details of collection

APPENDIX 3.3 – FOCUS GROUP SCRIPT[39] (GROUP 3)

To be read by the researcher at the beginning of the focus group:

Thank you for attending this focus group! This should take between 45 to 60 minutes, depending on your discussion. At the end you will receive a £10 Amazon voucher in compensation for your time, and to thank you for your participation.

This questionnaire is part of the *"Understanding the Flow of Information within User Generated Content in Virtual Education Learning Environments"* research. It is NOT part of your module and therefore: is optional, will NOT impact your grade in any way, and your answers will NOT be shared with the teaching staff on the module. Moreover, your identity will be anonymised at the end of the study, before any results are shown. The project has been inspected and approved by the *University of Sheffield Management School Ethics Committee* in accordance with the University of Sheffield ethics policy. You will now be given an information sheet regarding the purpose of this focus group and a consent form that you should fill-in if you decide to take part in the focus group.

(After everyone has given consent, turn on the recording devices and say the date).

Autumn Semester 2016-17 (Focus: rating scales)
Good day! Could I ask all of you to please introduce yourselves with your first/given name only? (Record all voices)

1. Could you tell me some of your fist impressions of PeerWise: good & bad.
2. PeerWise allowed you to sort and select questions in different ways: by author, by popularity (number of answers/comments), by difficulty, by number of likes, etc. Which was the one you find the most helpful and why?
3. How did you feel about being signed-in with your real names in PeerWise? // Do you think the amount of questions or comments you made would have been different if you had been assigned an anonymous ID?

[39] Note that, given the nature of focus groups, the questions shown here may have varied in order and/or content. Moreover, new questions might have surged while talking to participants, and are therefore not presented here.

4. Do you think being able to identify each other influenced in any way how you rated others?
5. How about the rating scale? Do you think yourself or others would have rated differently if instead of like/dislike you were given more options (e.g. a five or ten-point rating scale)?
6. When rating a question, do you think you rated it differently if the author was someone you knew? (Follow-up, optional: Let's say, if one of your friends wrote a bad question, would you rate it with a dislike, with a like, or would you rather not rate that question at all?)
7. Were you aware of your friend's questions? Did you followed them in PeerWise?
8. How did you feel about your name being (or not being) in the leaderboards?
9. Overall, do you think using PeerWise helped you learn? Why or why not?
10. Any other comments/questions?

Thank you for your time!

(Give vouchers & receipts)

APPENDIX 4.1 – NORMALITY TESTS FOR GROUPS 1 AND 2

Kolmogorov-Smirnov, Group 1

Tests of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Rating Score	.202	25927	.000

a. Lilliefors Significance Correction

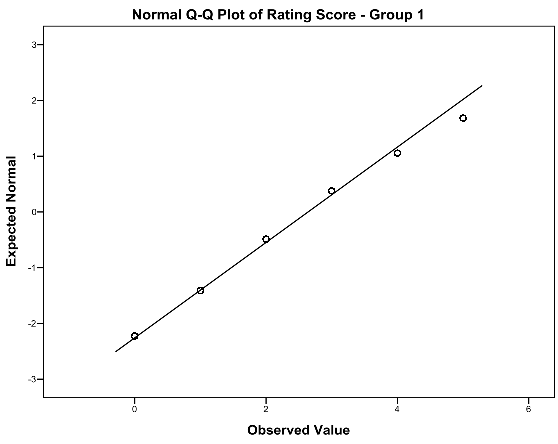
Kolmogorov-Smirnov, Group 2

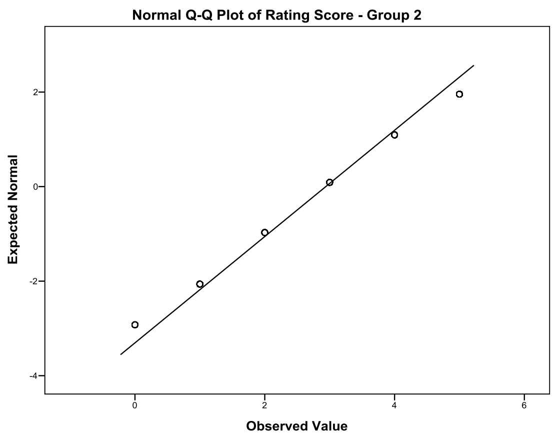
Tests of Normality

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Rating Score	.249	25976	.000

a. Lilliefors Significance Correction

Q-Q Plots, Groups 1 & 2





APPENDIX 4.2 – REGRESSION MODELS FOR GROUPS 1 AND 2

Group 1: R^2 = 5.7% variation; F = 194.44, sig < .001

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.238 [a]	.057	.056	1.135

a. Predictors: (Constant), Author's Distinct Badges, Qs Link, Qs Reference, Qs Explanation, Qs number of alternatives, Rating's date (by week), Qs number of characters, Author admin

ANOVA [a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2005.487	8	250.686	194.438	.000 [b]
	Residual	33415.613	25918	1.289		
	Total	35421.099	25926			

a. Dependent Variable: Rating Score

b. Predictors: (Constant), Author's Distinct Badges, Qs Link, Qs Reference, Qs Explanation, Qs number of alternatives, Rating's date (by week), Qs number of characters, Author admin

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1.085	.154		7.051	.000	.783	1.386
	Qs Explanation	.363	.136	.016	2.665	.008	.096	.630
	Qs number of alternatives	.035	.011	.020	3.197	.001	.014	.057
	Qs Reference	.254	.021	.075	12.224	.000	.213	.294
	Qs Link	.162	.030	.033	5.361	.000	.102	.221
	Author admin	.909	.064	.100	14.147	.000	.783	1.035
	Qs number of characters	.001	.000	.146	23.253	.000	.001	.001
	Rating's date (by week)	-.031	.005	-.040	-6.284	.000	-.040	-.021
	Author's Distinct Badges	.053	.003	.147	21.067	.000	.048	.058

a. Dependent Variable: Rating Score

Group 2: R^2 = 7.1% variation; F = 249.76.03, sig < .001

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1.336	.091		14.749	.000	1.158	1.513
	Qs Explanation	.283	.081	.021	3.496	.000	.124	.441
	Qs number of alternatives	.052	.005	.065	10.150	.000	.042	.062
	Qs Reference	.175	.015	.072	11.839	.000	.146	.203
	Qs Link	.093	.011	.052	8.195	.000	.071	.115
	Author admin	.544	.047	.083	11.535	.000	.452	.636
	Qs number of characters	.001	.000	.149	22.776	.000	.000	.001
	Rating's date (by week)	.085	.004	.150	23.591	.000	.078	.092
	Author's Distinct Badges	.024	.002	.106	15.276	.000	.021	.027

a. Dependent Variable: Rating Score

APPENDIX 4.3 – CORRELATION & REGRESSION FOR BEING IDENTIFIABLE & WEBSITE'S PERCEPTIONS (GROUP 2)

Spearman's rho coefficient of correlation

Correlations [a]

			PeerWise's perception	Identifiable usernames' perception
Spearman's rho	PeerWise's perception	Correlation Coefficient	1.000	.128
		Sig. (2-tailed)	.	.099
		N	167	167
	Identifiable usernames' perception	Correlation Coefficient	.128	1.000
		Sig. (2-tailed)	.099	.
		N	167	186

a. Group = Group 2

Ordinal regression

Model Fitting Information [a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	62.231			
Final	57.184	5.048	4	.282

Link function: Logit.

Pseudo R-Square [a]

Cox and Snell	.030
Nagelkerke	.034
McFadden	.015

Link function: Logit.

APPENDIX 5.1 – REGRESSIONS AGGREGATED AT THE QUESTION LEVEL (GROUPS 2 AND 3)

Group 2

Model Summary [a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
2	.447 [c]	.200	.197	.47365

a. Group = Group 2

c. Predictors: (Constant), Number of ratings, Question's Explanation, Author's Gender (c), Question's number of alternatives, Friends (Prop.), Author's Nationality (c), Question's Link, Question's Reference, Author's Distinct Badges, Question's number of characters

ANOVA [a,b]

Model		Sum of Squares	df	Mean Square	F	Sig.
2	Regression	134.499	10	13.450	59.952	.000 [d]
	Residual	537.980	2398	.224		
	Total	672.479	2408			

a. Group = Group 2

b. Dependent Variable: Average Rating Score

Coefficients [a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
2	(Constant)	2.035	.162		12.587	.000
	Friends (Prop.)	.785	.067	.219	11.685	.000
	Author's Distinct Badges	.017	.003	.115	5.978	.000
	Author's Gender (c)	-.124	.020	-.116	-6.314	.000
	Author's Nationality (c)	-.032	.012	-.049	-2.619	.009
	Question's Explanation	.416	.150	.051	2.766	.006
	Question's number of alternatives	.033	.009	.067	3.514	.000
	Question's Reference	.105	.025	.078	4.113	.000
	Question's Link	.138	.020	.130	6.871	.000
	Question's number of characters	.000	.000	.206	10.545	.000
	Number of ratings	-.006	.001	-.106	-5.383	.000

a. Group = Group 2

b. Dependent Variable: Average Rating Score

Group 3

Model Summary [a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
6	.161 [g]	.026	.023	.74213

a. Group = Group 3

g. Predictors: (Constant), Question's Link, Question's Explanation, Friends (Prop.), Question's number of characters, Author's Nationality (c), Author's Distinct Badges

ANOVA [a,b]

Model		Sum of Squares	df	Mean Square	F	Sig.
6	Regression	28.686	6	4.781	8.681	.000 [h]
	Residual	1079.490	1960	.551		
	Total	1108.176	1966			

a. Group = Group 3

b. Dependent Variable: Average Rating Score

Coefficients [a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
6	(Constant)	4.206	.115		36.710	.000
	Friends (Prop.)	.255	.081	.073	3.130	.002
	Author's Distinct Badges	.007	.004	.044	1.927	.054
	Author's Nationality (c)	-.037	.021	-.039	-1.714	.087
	Question's Explanation	.251	.100	.057	2.517	.012
	Question's Link	.134	.035	.066	3.832	.000
	Question's number of characters	.000	.000	.047	2.043	.041

a. Group = Group 3

b. Dependent Variable: Average Rating Score

APPENDIX 5.2 – THREE WAY CHI-SQUARE: THE EFFECT OF RATING SCALES ON FRIENDSHIPS, PER GROUP (GROUPS 2 AND 3)

3-way chi square: *Groups (2 and 3), Friendships (absent or present) & Rating (Like and Dislike)*

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Group * Friendships (undirected) * Rating (Dichotomous)	40177	98.1%	796	1.9%	40973	100.0%

Group * Friendships (undirected) * Rating (Dichotomous) Crosstabulation

Count

Rating (Dichotomous)			Friendships (undirected): No relationship	Friendships (undirected): Some relationship	Total
Dislike	Group	Group 2	7333	193	7526
		Group 3	859	45	904
	Total		8192	238	8430
Like	Group	Group 2	16111	1854	17965
		Group 3	11864	1918	13782
	Total		27975	3772	31747
Total	Group	Group 2	23444	2047	25491
		Group 3	12723	1963	14686
	Total		36167	4010	40177

Chi-Square Tests

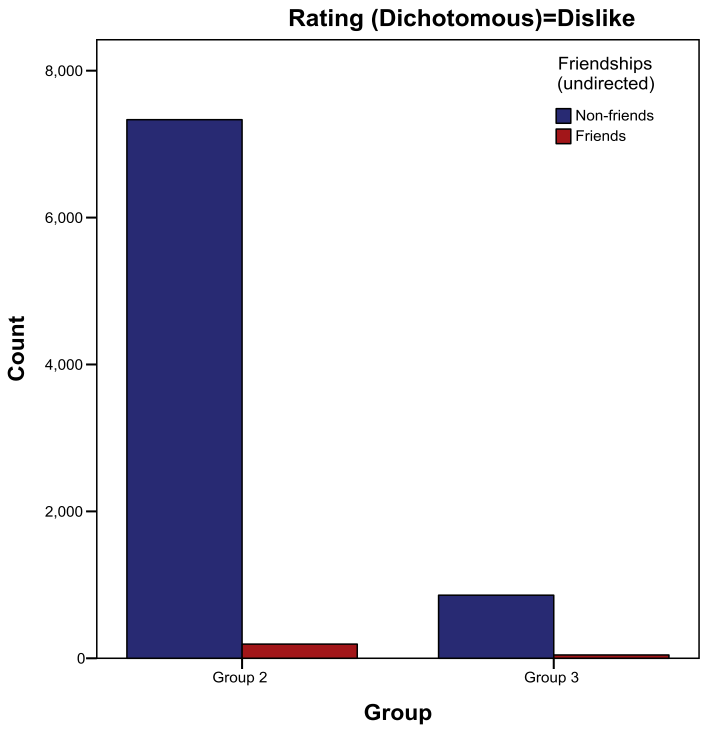
Rating (Dichotomous)		Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Dislike	Pearson Chi-Square	17.134 [c]	1	.000		
	Continuity Correction [b]	16.266	1	.000		
	Likelihood Ratio	14.414	1	.000		
	Fisher's Exact Test				.000	.000
	Linear-by-Linear Association	17.132	1	.000		
	N of Valid Cases	8430				
Like	Pearson Chi-Square	96.359 [d]	1	.000		
	Continuity Correction [b]	96.016	1	.000		
	Likelihood Ratio	95.580	1	.000		
	Fisher's Exact Test				.000	.000
	Linear-by-Linear Association	96.356	1	.000		
	N of Valid Cases	31747				
Total	Pearson Chi-Square	295.306 [a]	1	.000		
	Continuity Correction [b]	294.713	1	.000		
	Likelihood Ratio	286.126	1	.000		
	Fisher's Exact Test				.000	.000
	Linear-by-Linear Association	295.299	1	.000		
	N of Valid Cases	40177				

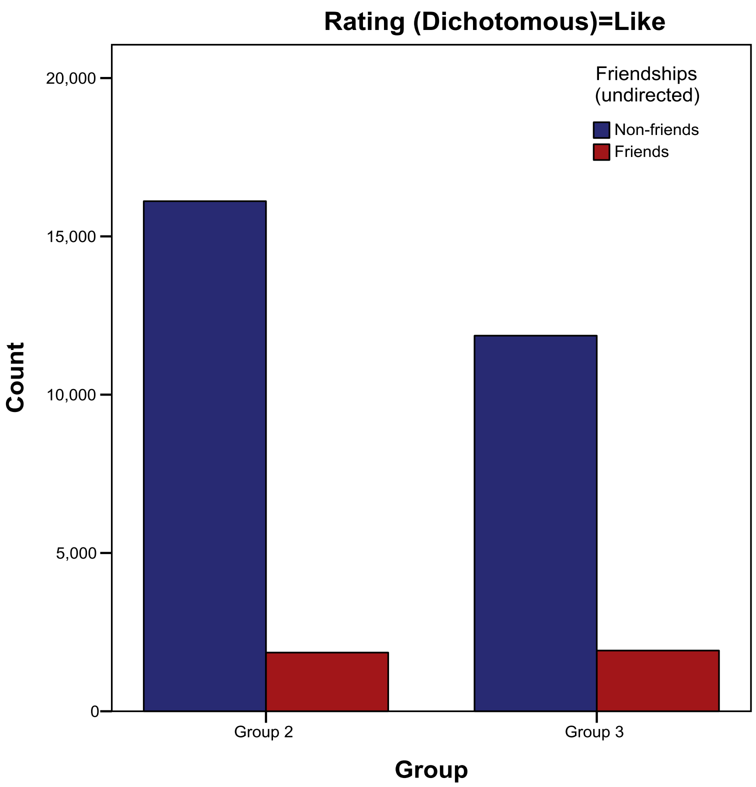
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 1465.79.

b. Computed only for a 2x2 table

c. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 25.52.

d. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 1637.50.


Rating (Dichotomous)=Dislike
Friendships (undirected)
Non-friends
Friends
Count
8,000
6,000
4,000
2,000
0
Group 2
Group 3
Group


Rating (Dichotomous)=Like
Friendships (undirected)
Non-friends
Friends
Count
20,000
15,000
10,000
5,000
0
Group 2
Group 3
Group

APPENDIX 5.3 – CORRELATION & REGRESSION FOR RATING SCALE & WEBSITE'S PERCEPTIONS (GROUPS 2 AND 3)

Group 2

Correlations [a]

			PeerWise's perception	Rating scale's perception
Spearman's rho	PeerWise's perception	Correlation Coefficient	1.000	.374**
		Sig. (2-tailed)	.	.000
		N	167	167
	Rating scale's perception	Correlation Coefficient	.374**	1.000
		Sig. (2-tailed)	.000	.
		N	167	187

**. Correlation is significant at the 0.01 level (2-tailed).

a. Group = Group 2

Model Fitting Information [a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	88.699			
Final	48.262	40.437	4	.000

Link function: Logit.

a. Group = Group 2

Pseudo R-Square [a]

Cox and Snell	.215
Nagelkerke	.246
McFadden	.117

Link function: Logit.

a. Group = Group 2

Group 3

Correlations [a]

			PeerWise's perception	Rating scale's perception
Spearman's rho	PeerWise's perception	Correlation Coefficient	1.000	.329**
		Sig. (2-tailed)	.	.000
		N	200	200
	Rating scale's perception	Correlation Coefficient	.329**	1.000
		Sig. (2-tailed)	.000	.
		N	200	206

**. Correlation is significant at the 0.01 level (2-tailed).

a. Group = Group 3

Model Fitting Information [a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	114.149			
Final	66.082	48.067	4	.000

Link function: Logit.

a. Group = Group 3

Pseudo R-Square [a]

Cox and Snell	.214
Nagelkerke	.241
McFadden	.110

Link function: Logit.

a. Group = Group 3

APPENDIX 6.1 – ANOVAS FOR STRENGTHS OF TIES, GROUPS 2 AND 3

ANOVA for the strength of ties for friendships in Group 2

Case Processing Summary [a]

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Rating (Scale 0 to 5) * Strength of ties (friendships)	25445	98.0%	531	2.0%	25976	100.0%

a. Group = Group 2

Report [a]

Rating (Scale 0 to 5)

Strength of ties (friendships)	Mean	Std. Error of Mean	N	Std. Deviation
No relationship	2.88	.006	23444	.865
Weak	3.32	.057	264	.926
Medium	3.59	.037	560	.876
Strong	3.73	.025	1177	.851
Total	2.94	.006	25445	.890

a. Group = Group 2

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Strength of ties (friendships)	Between Groups	(Combined)	1091.160	3	363.720	485.355	.000
		Linearity	1088.774	1	1088.774	1452.880	.000
		Deviation from Linearity	2.386	2	1.193	1.592	.203
	Within Groups		19065.231	25441	.749		
	Total		20156.391	25444			

a. Group = Group 2

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Strength of ties (friendships)	.232	.054	.233	.054

a. Group = Group 2

ANOVA for the strength of ties for friendships in Group 3

Case Processing Summary [a]

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Rating (Scale 0 to 5) * Strength of ties (friendships)	14616	97.5%	381	2.5%	14997	100.0%

a. Group = Group 3

Report [a]

Rating (Scale 0 to 5)

Strength of ties (friendships)	Mean	Std. Error of Mean	N	Std. Deviation
No relationship	4.66	.011	12723	1.255
Weak	4.83	.052	300	.899
Medium	4.83	.038	579	.914
Strong	4.94	.017	1014	.541
Total	4.69	.010	14616	1.203

a. Group = Group 3

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Strength of ties (friendships)	Between Groups	(Combined)	90.502	3	30.167	20.945	.000
		Linearity	89.228	1	89.228	61.950	.000
		Deviation from Linearity	1.274	2	.637	.442	.643
	Within Groups		21045.949	14612	1.440		
	Total		21136.451	14615			

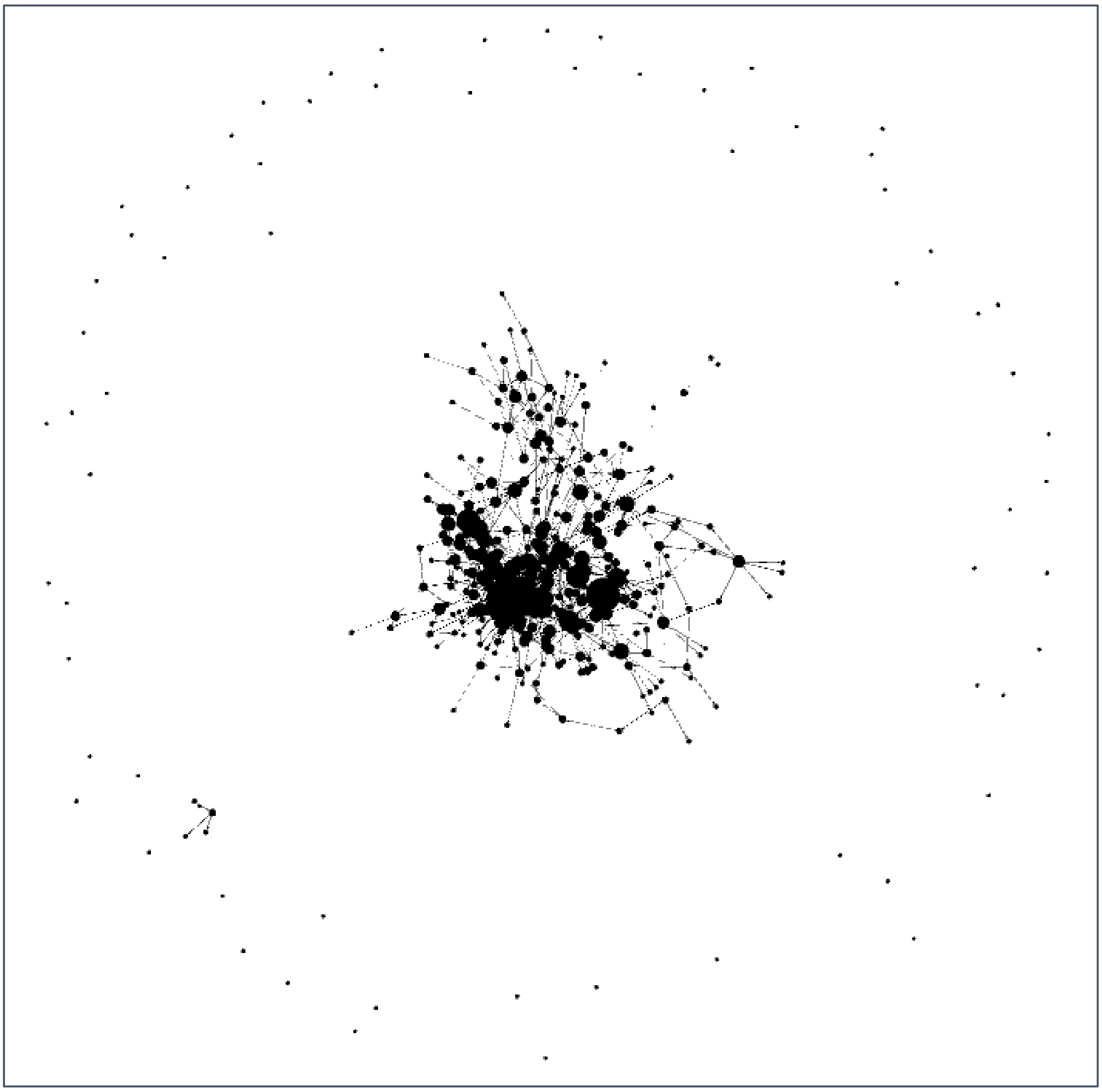
a. Group = Group 3

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Strength of ties (friendships)	.065	.004	.065	.004

a. Group = Group 3

APPENDIX 6.2 – ORIGINAL NETWORK IMAGE (GROUP 2)



APPENDIX 6.3 – COMPLETE SET OF GEODESIC DISTANCES AND ANOVAS FROM FRIENDSHIPS, GROUPS 2 AND 3

Geodesic Distances from friendships – Groups 2 and 3

	Group 2			Group 3		
Geo Dist	**No. Ratings**	**Pct.**	**Ave. Rating**	**No. Ratings**	**Pct.**	**Ave. Rating**
GeoD_1	1,743	7%	3.68	1,606	11%	4.90
GeoD_2	1,938	8%	3.38	2,292	16%	4.82
GeoD_3	2,504	10%	2.99	3,235	22%	4.68
GeoD_4	3,471	14%	2.87	3,958	27%	4.62
GeoD_5	3,646	14%	2.77	2,152	15%	4.61
GeoD_6	2,977	12%	2.76	722	5%	4.57
GeoD_7	1,765	7%	2.79	150	1%	4.63
GeoD_8	883	3%	2.79	30	0%	4.50
GeoD_9	311	1%	2.92	2	0%	5.00
GeoD_10	117	0%	2.77	-	-	-
GeoD_11	84	0%	3.18	-	-	-
GeoD_12	22	0%	2.77	-	-	-
GeoD_13	12	0%	3.17	-	-	-
GeoD_14	5	0%	3.20	-	-	-
GeoD_15	1	0%	2.00	-	-	-
GeoD_16	0	0%	2.00	-	-	-
GeoD_17	1	0%	-	-	-	-
GeoD_Inf [40]	119	0%	2.65	-	-	-
GeoD_NA[41]	5,892	23%	2.87	539	4%	4.65
TOTAL*	***25,491***	***100%***	***2.94***	***14,686***	***100%***	***4.69***

[40] ***GeoD_Inf*** is when two people do appear in the network but there is not a geodesic path that can connect them.
[41] ***GeoD_NA*** is when there was a rating between people who did not appear on the friendships network, i.e. when someone did not answered the survey and no one named them as a friend/acquaintance.

ANOVA for geodesic distances of friendships in Group 2

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Geodesic Distance Friends	Between Groups	(Combined)	1615.264	9	179.474	254.547	.000
		Linearity	1047.700	1	1047.700	1485.948	.000
		Deviation from Linearity	567.564	8	70.946	100.622	.000
	Within Groups		13639.617	19345	.705		
	Total		15254.882	19354			

a. Group = Group 2

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Geodesic Distance Friends	-.262	.069	.325	.106

a. Group = Group 2

ANOVA for geodesic distances of friendships in Group 3

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Geodesic Distance Friends	Between Groups	(Combined)	150.847	6	25.141	17.642	.000
		Linearity	129.700	1	129.700	91.013	.000
		Deviation from Linearity	21.147	5	4.229	2.968	.011
	Within Groups		20105.044	14108	1.425		
	Total		20255.891	14114			

a. Group = Group 3

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Geodesic Distance Friends	-.080	.006	.086	.007

a. Group = Group 3

APPENDIX 6.4 – COMPARISON OF SNA METHODS, GROUP 2

The following regressions were conducted at the question-level: 1) Friendships, 2) Strengths of ties, 3) Clusters, 4) Geodesic distances.

1) Friendships (proportion of friendships: absent/present, Group 2)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.224 [a]	.050	.050	.868

a. Predictors: (Constant), Friendships (undirected)

ANOVA [a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1012.482	1	1012.482	1345.030	.000 [b]
	Residual	19187.046	25489	.753		
	Total	20199.527	25490			

a. Dependent Variable: Rating (Scale 0 to 5)

b. Predictors: (Constant), Friendships (undirected)

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.881	.006		508.467	.000
	Friendships (undirected)	.733	.020	.224	36.675	.000

a. Dependent Variable: Rating (Scale 0 to 5)

2) Strength of ties (average strength of ties, Group 2)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.232 [a]	.054	.054	.866

a. Predictors: (Constant), Strength of ties (friendships)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1088.774	1	1088.774	1452.813	.000 [b]
	Residual	19067.617	25443	.749		
	Total	20156.391	25444			

a. Dependent Variable: Rating (Scale 0 to 5)

b. Predictors: (Constant), Strength of ties (friendships)

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.882	.006		510.556	.000
	Strength of ties (friendships)	.217	.006	.232	38.116	.000

a. Dependent Variable: Rating (Scale 0 to 5)

3) Clusters (proportion of same cluster: same/different cluster, Group 2)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.243 [a]	.059	.059	.864

a. Predictors: (Constant), Cluster

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1189.273	1	1189.273	1594.580	.000 [b]
	Residual	19010.255	25489	.746		
	Total	20199.527	25490			

a. Dependent Variable: Rating (Scale 0 to 5)

b. Predictors: (Constant), Cluster

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.846	.006		482.313	.000
	Cluster	.590	.015	.243	39.932	.000

a. Dependent Variable: Rating (Scale 0 to 5)

4) Geodesic distances (average geodesic distance, Group 2)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.262 [a]	.069	.069	.857

a. Predictors: (Constant), Geodesic Distance Friends

ANOVA [a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1047.700	1	1047.700	1427.175	.000 [b]
	Residual	14207.181	19353	.734		
	Total	15254.882	19354			

a. Dependent Variable: Rating (Scale 0 to 5)

b. Predictors: (Constant), Geodesic Distance Friends

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.477	.015		232.876	.000
	Geodesic Distance Friends	-.115	.003	-.262	-37.778	.000

a. Dependent Variable: Rating (Scale 0 to 5)

APPENDIX 6.5 – MULTILEVEL MODELLING, GROUP 2

1) Unconditional Model (no predictors)

Information Criteria [a]

-2 Log Likelihood	63220.458
Akaike's Information Criterion (AIC)	63228.458
Hurvich and Tsai's Criterion (AICC)	63228.459
Bozdogan's Criterion (CAIC)	63265.249
Schwarz's Bayesian Criterion (BIC)	63261.249

-2LL (unconditional)

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: Score given by the rater to the author.

Type III Tests of Fixed Effects [a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	594.194	14199.233	.000

a. Dependent Variable: Score given by the rater to the author.

Estimates of Fixed Effects [a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	2.995178	.025136	594.194	119.161	.000	2.945813	3.044544

a. Dependent Variable: Score given by the rater to the author.

Estimates of Covariance Parameters [a]

	Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
1st	Residual		.578863	.005065	114.296	.000	.569022	.588875
2nd	Intercept [subject = rater]	Variance	.152831	.012403	12.322	.000	.130357	.179181
	Intercept [subject = author]	Variance	.061446	.005420	11.338	.000	.051691	.073042

a. Dependent Variable: Score given by the rater to the author.

$$ICC(1) = \frac{2nd\ level\ variance}{1st + 2nd\ level\ variance} = \frac{0.1528 + 0.0614}{0.5789 + 0.1528 + 0.0614} = 0.270163 = 27.0\%$$

2) Model with ‘same_cluster’ as a level-2 predictor

Information Criteria [a]

-2 Log Likelihood	61798.677
Akaike's Information Criterion (AIC)	61808.677
Hurvich and Tsai's Criterion (AICC)	61808.679
Bozdogan's Criterion (CAIC)	61854.666
Schwarz's Bayesian Criterion (BIC)	61849.666

-2LL
(with predictor)

The information criteria are displayed in smaller-is-better form.

a. Dependent Variable: Score given by the rater to the author.

Simple chi square test - Excel sheet (Stride, 2016)

	Model Deviance (-2LL)	Number of new variables added to make more complex model	p for difference test
Simpler Model	63,220.46		
More complex model	61,798.67		
Difference Test	1,421.79	1	0.0000000000000000000000000000000

Type III Tests of Fixed Effects [a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	648.121	16454.937	.000
same_cluster	1	26816.612	1460.527	.000

a. Dependent Variable: Score given by the rater to the author.

Estimates of Fixed Effects [a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	3.445668	.027005	910.389	127.594	.000	3.392669	3.498667
[same_cluster=0]	-.543618	.014225	26816.612	-38.217	.000	-.571499	-.515737
[same_cluster=1]	0[b]	0	.	.	.	.	.

a. Dependent Variable: Score given by the rater to the author.

Estimates of Covariance Parameters [a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		.549038	.004804	114.296	.000	.539703	.558534
Intercept [subject = rater]	Variance	.139429	.011356	12.278	.000	.118857	.163562
Intercept [subject = author]	Variance	.060485	.005292	11.430	.000	.050953	.071799

a. Dependent Variable: Score given by the rater to the author.

APPENDIX 6.6 – COMPLETE SET OF GEODESIC DISTANCES AND ANOVAS FROM FOLLOWERS (GROUPS 1, 2 AND 3)

Geodesic Distances from followers – Groups 1, 2 and 3

	Group 1			Group 2			Group 3		
Geo Dist	**No. Rtngs**	**Pct.**	**Ave Rtg**	**No. Rtngs**	**Pct.**	**Ave Rtg**	**No. Rtngs**	**Pct.**	**Ave Rtg**
GeoD_1	5,741	23%	3.58	5,466	21%	3.68	3,847	26%	4.92
GeoD_2	2,662	10%	2.45	2,700	11%	2.98	1,598	11%	4.76
GeoD_3	5,325	21%	2.35	5,070	20%	2.73	2,423	16%	4.58
GeoD_4	6,447	25%	2.35	5,466	21%	2.68	2,270	15%	4.58
GeoD_5	2,941	12%	2.43	2,948	12%	2.66	1,135	8%	4.54
GeoD_6	741	3%	2.39	893	4%	2.68	362	2%	4.56
GeoD_7	117	0%	2.41	201	1%	2.61	165	1%	4.48
GeoD_8	3	0%	2.33	14	0%	2.14	165	1%	4.18
GeoD_9	-	-	-	4	0%	2.75	95	1%	4.21
GeoD_10	-	-	-	-	-	-	18	0%	4.72
GeoD_11	-	-	-	-	-	-	6	0%	4.17
GeoD_12	-	-	-	-	-	-	-	-	-
GeoD_Inf [42]	-	-	-	-	-	-	865	6%	4.80
GeoD_NA[43]	1,513	6%	2.17	2,729	11%	2.75	1,737	12%	4.59
TOTAL*	***25,490***	***100%***	***2.64***	***25,491***	***100%***	***2.94***	***14,686***	***100%***	***4.69***

ANOVA for geodesic distances of followers in Group 2

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Geodesic Distance Followers	Between Groups	(Combined)	3887.836	6	647.973	1020.775	.000
		Linearity	2822.825	1	2822.825	4446.901	.000
		Deviation from Linearity	1065.011	5	213.002	335.550	.000
	Within Groups		14438.814	22746	.635		
	Total		18326.650	22752			

a. Group = Group 2

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Geodesic Distance Followers	-.392	.154	.461	.212

a. Group = Group 2

[42] ***GeoD_Inf*** is when two people do appear in the network but there is not a geodesic path that can connect them.
[43] ***GeoD_NA*** is when there was a rating between people who did not appear on the friendships network, i.e. when someone did not answered the survey and no one named them as a friend/acquaintance.

ANOVA for geodesic distances of 'followers' in Group 3

ANOVA Table [a]

			Sum of Squares	df	Mean Square	F	Sig.
Rating (Scale 0 to 5) * Geodesic Distance Followers	Between Groups	(Combined)	355.718	7	50.817	37.184	.000
		Linearity	300.122	1	300.122	219.605	.000
		Deviation from Linearity	55.596	6	9.266	6.780	.000
	Within Groups		16341.002	11957	1.367		
	Total		16696.720	11964			

a. Group = Group 3

Measures of Association [a]

	R	R Squared	Eta	Eta Squared
Rating (Scale 0 to 5) * Geodesic Distance Followers	-.134	.018	.146	.021

a. Group = Group 3

ANOVA for geodesic distances of 'followers' in Group 1 (i.e. *Inferred friendships)*

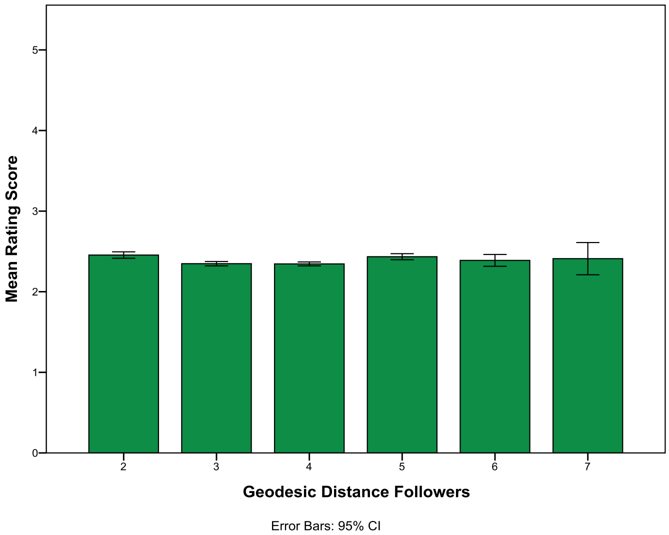
ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Rating Score * Geodesic Distance Followers	Between Groups	(Combined)	6299.421	6	1049.903	935.557	.000
		Linearity	3773.730	1	3773.730	3362.727	.000
		Deviation from Linearity	2525.691	5	505.138	450.123	.000
	Within Groups		26896.318	23967	1.122		
	Total		33195.739	23973			

Measures of Association

	R	R Squared	Eta	Eta Squared
Rating Score * Geodesic Distance Followers	-.337	.114	.436	.190

ANOVA and R^2 for geodesics of followers in Group 1, removing geodesic distance = 1



ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Rating Score * Geodesic Distance Followers	Between Groups	(Combined)	37.121	5	7.424	7.213	.000
		Linearity	.132	1	.132	.128	.720
		Deviation from Linearity	36.989	4	9.247	8.984	.000
	Within Groups		18760.408	18227	1.029		
	Total		18797.529	18232			

Measures of Association

	R	R Squared	Eta	Eta Squared
Rating Score * Geodesic Distance Followers	-.003	.000	.044	.002

APPENDIX 7.1 – ANOVAS FOR SEQUENCE MEASURES (GROUPS 1, 2 AND 3)

Original ratings

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Sequence Length	Between Groups	25908.549	2	12954.274	170.944	.000
	Within Groups	491589.119	6487	75.781		
	Total	517497.667	6489			
Number of transitions (Original)	Between Groups	54575.001	2	27287.501	994.860	.000
	Within Groups	177928.619	6487	27.428		
	Total	232503.620	6489			
Within entropy (Original)	Between Groups	310.756	2	155.378	4714.615	.000
	Within Groups	213.790	6487	.033		
	Total	524.545	6489			
Complexity Index (Original)	Between Groups	338.556	2	169.278	4607.044	.000
	Within Groups	238.354	6487	.037		
	Total	576.911	6489			

First conversion: [0,1,2 = Dislike] & [3,4,5 = Like]

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Number of transitions (Converted 3&3)	Between Groups	16523.384	2	8261.692	610.072	.000
	Within Groups	87848.024	6487	13.542		
	Total	104371.408	6489			
Within entropy (Converted 3&3)	Between Groups	362.939	2	181.470	1536.338	.000
	Within Groups	766.233	6487	.118		
	Total	1129.173	6489			
Complexity Index (Converted 3&3)	Between Groups	184.945	2	92.473	1259.654	.000
	Within Groups	476.218	6487	.073		
	Total	661.163	6489			

Second conversion: [0,1 = Dislike] & [2,3,4,5 = Like]

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Number of transitions (Converted 2&4)	Between Groups	3497.852	2	1748.926	363.520	.000
	Within Groups	31209.523	6487	4.811		
	Total	34707.374	6489			
Within entropy (Converted 2&4)	Between Groups	95.941	2	47.971	529.591	.000
	Within Groups	587.595	6487	.091		
	Total	683.536	6489			
Complexity Index (Converted 2&4)	Between Groups	45.848	2	22.924	444.118	.000
	Within Groups	334.837	6487	.052		
	Total	380.685	6489			

APPENDIX 7.2 – CORRELATIONS AND ORDINAL REGRESSIONS BETWEEN ALL ATTITUDES (GROUPS 2 AND 3)

Group 2

Correlations [a]

			PeerWise's perception	Rating scale's perception	Identifiable usernames' perception
Spearman's rho	PeerWise's perception	Correlation Coefficient	1.000	.374**	.128
		Sig. (2-tailed)	.	.000	.099
		N	167	167	167
	Rating scale's perception	Correlation Coefficient	.374**	1.000	.204**
		Sig. (2-tailed)	.000	.	.005
		N	167	187	186
	Identifiable usernames' perception	Correlation Coefficient	.128	.204**	1.000
		Sig. (2-tailed)	.099	.005	.
		N	167	186	186

**. Correlation is significant at the 0.01 level (2-tailed).

a. Group = Group 3

Model Fitting Information [a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	157.704			
Final	116.183	41.520	8	.000

Link function: Logit.

a. Group = Group 2

Pseudo R-Square [a]

Cox and Snell	.220
Nagelkerke	.252
McFadden	.121

Link function: Logit.

a. Group = Group 2

Group 3

Correlations [a]

			PeerWise's perception	Rating scale's perception	Identifiable usernames' perception
Spearman's rho	PeerWise's perception	Correlation Coefficient	1.000	.374**	.128
		Sig. (2-tailed)	.	.000	.099
		N	167	167	167
	Rating scale's perception	Correlation Coefficient	.374**	1.000	.204**
		Sig. (2-tailed)	.000	.	.005
		N	167	187	186
	Identifiable usernames' perception	Correlation Coefficient	.128	.204**	1.000
		Sig. (2-tailed)	.099	.005	.
		N	167	186	186

**. Correlation is significant at the 0.01 level (2-tailed).

a. Group = Group 3

Model Fitting Information [a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	211.795			
Final	151.147	60.648	8	.000

Link function: Logit.

a. Group = Group 3

Pseudo R-Square [a]

Cox and Snell	.262
Nagelkerke	.295
McFadden	.139

Link function: Logit.

a. Group = Group 3