EMOTIONS AND RECOMMENDER SYSTEMS: A SOCIAL NETWORK APPROACH

Carlos Manuel Moita de Figueiredo

Dissertação apresentada na Faculdade de Engenharia da

Universidade do Porto, para obtenção do grau de Doutor em Media Digitais

Orientador: José Azevedo

Co-orientador: Eduarda Mendes Rodrigues

2014

Doutoramento financiado por uma bolsa de doutoramento com a referência
SFRH / BD / 51305 / 2010

To all those who helped me along the way.

To the smile of my nephew and nieces.

ACKNOWLEDGMENTS

These last four years of work have been immensely challenging and positive. This journey started with a question: “can we recommend content to online users in the context of their browsing activity and trigger richer ways of interpreting it through new insights?” This question arose while I was managing a startup company named "Edit on Web" during the period of 2003 - 2010. This amazing project of networking between people and content centered on contributing new ways of publishing, browsing and accessing information online for users and by users, benefiting from new insights by means of recommendation, ended up to be one of the triggers to start this thesis. In fact, because the seven years of the company's existence were not enough to find the right answers, I decided to look for it by way of this dissertation.

However, this journey wouldn’t be possible without the knowledge, support and friendship of so many good people and family.

I would like to express my sincere gratitude to these people, without whom this thesis wouldn’t have reached this point and my PhD journey wouldn’t have been as exciting and enjoyable as it has been.

I am especially grateful for the mentorship of my primary supervisor Dr. José Azevedo, who since the first moment has been present and ready to guide me along the path. Without our first meeting talking about the "Edit on Web" at Porto University (UP), and his vision, this journey couldn't have started. Thanks also go to Professor Eurico Carrapatoso, who welcomed me and advised me in the early stages of this journey, with his unwavering and contagious smile, and to my second supervisor Dr. Eduarda Mendes Rodrigues, who provided comments at times when it was crucial.

Another big thank you goes to Dr. Luis Francisco-Revilla and Dr. Wenhong Chen, who were wonderful supervisors during my three semesters at the University of Texas in Austin (UT). It is difficult to imagine the end of this dissertation without their guidance, talks, comments, networking and above all, questions and support.

Several professors at UT have been a 'light' to my path during this journey, so thank you all, and particularly to Derek Lackaff, Wenhong Chen and Matt Lease.

Without them it would not have been possible to find the main building blocks of this dissertation.

Thank you to Christiano Prado, João Beira, Ryan Chen, Jeremy Tsuy, and so many others for being such great hosts during the time I spent in Austin. Special thanks to Filipa Ribeiro who offered me unsurpassed support and to Sara Cavaco for the wonderful help with the data and statistical work. Thank you to Susana Silva, Carla C. Martins, Anabela Costa, Nuno Natividade, Kelly Mendonça and Paddy Greenleaf for the friendship and so much advice, chats and help.

My dissertation research was made possible by the UT – Austin | Portugal Doctoral Program. It is impossible to imagine it happening without the vision of professor Pimenta Alves and Sharon Strover and the efficiency of Karen Gustafson. Thank you all and to the FEUP and RTF team dedicated to help on this UT-PT Program.

Finally, I am grateful to my wonderful parents, brother and sister. You have always been an incredible support and not least over the past 4 years. Thank you for your love and your commitment to me through all the ups and downs. You show me why there is no aim, without ties. And my appreciation to Coronel Costeira, for his amazing work and solidarity with so many, making explicit the implicit, who helped me to find focus, learning and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

RESUMO

Esta dissertação contribui para os campos científicos sobre Sistemas de Recomendação e Análise de Redes Sociais, mas também para os estudos em Sociopsicologia aplicados a Sistemas de Media Digitais. As empresas de *media* digital estão a usar dados de redes sociais para personalizar serviços baseados na Web (por exemplo, busca e recomendação) para servir e envolver a sua audiência de forma mais eficaz e relevante. Porém, esta prática está a diminuir a diversidade de pontos de vista na comunidade de utilizadores Web dada a falta de novidade nos resultados entregues. Assim, o uso actual de dados sociais baseados em relações estabelecidas por efeitos endógenos (ou seja, homofilia) e amizade ou proximidade social (ou seja, laços fortes) cria um efeito de Câmara de Eco Social que aprisiona as pessoas dentro de bolhas sociais de informação. Por consequência, em vez de inovação, há uma redução de qualidade nos serviços prestados por sistemas de recomendação, e assim, um baixo nível de satisfação dos seus utilizadores.

Reconhecendo-se as desvantagens da utilização de dados de redes sociais, mas também a sua riqueza, este trabalho propõe-se a encontrar uma solução para a construção de um fluxo de informações e recomendações baseadas na novidade através de dados sociais. Três estudos empíricos apoiados numa abordagem interdisciplinar entre Análise de Redes Sociais e Psicologia e Neurociências, pesquisam que factores estruturais e atributos pessoais contribuem para a percepção de novidade. Estes estudos consideram em conjunto o estudo dos laços sociais e das semelhanças entre uma população de estudantes, bem como a sua resposta emocional à selecção de conteúdo em uma rede social. O primeiro estudo, que propõe um método alternativo para o estudo de pontes de rede e que se centra na análise da percepção da novidade pelos receptores de informação, apoia-se na hipótese de que a surpresa é um *proxy* da novidade, pelo que, os factores de ponte, ou seja, força do laço e buracos estruturais, podem ser analisados como preditores da resposta surpresa. O segundo estudo empírico baseia-se na constatação de que a selecção de conteúdos pelos receptores é mais dependente da reacção emocional do receptor, do que de factores associados à popularidade dos remetentes, ou à relação de amizade, proximidade, entre emissor e receptor. O último estudo analisa a distância cognitiva óptima, entre emissor e receptor, medido a partir de atributos pessoais que em conjunto com factores de ponte predizem a resposta surpresa do receptor.

Os resultados mostram que o desempenho dos sistemas de recomendação baseados em redes sociais pode ser melhorado através da entrega de recomendações novos e surpreendentes com base na previsão e não na aleatoriedade, o que evita o efeito de Câmara de Eco Social. Esta dissertação chama a atenção para o fato de que os dados sociais podem ser usados para aumentar as distâncias cognitivas entre os utilizadores da Web, o que permite lidar com um conjunto de novas ameaças (por exemplo, ao nível da democracia / tolerância, conformidade, cognição, e da inovação "fluffy"), que têm sido impostas por alguns algoritmos Web.

ABSTRACT

This dissertation contributes to the scientific fields of Recommender Systems and Social Network Analysis, but also to Social-psychological studies applied to Digital Media Systems. Digital media entrepreneurs are using data from social networks to personalize Web-based services (e.g., searching and recommendation) to engage their publics in more effective and striking ways. However, this practice is narrowing the diversity of viewpoints in the Web community because of the lack of access to novelty. I claim that the use of the current type of social data, based on relationships set by endogenous effects (i.e., homophily) and friendship or social proximity (i.e., strong ties) creates a Social Echo Chamber Effect that traps people inside social bubbles of information. Consequently, instead of innovation, there is a reduction of quality in the services provided by recommender systems, and so, a lower level of user’s satisfaction.

Acknowledging the drawbacks using data from social networks, but also its richness, this work proposes to find a solution to construct a flow of information and recommendations based on novelty through social data. Three empirical studies supported by an interdisciplinary approach between Social Network Analysis, Psychology and Neuroscience, examine which structural factors and personal attributes contribute to novelty perception. These studies consider in tandem the study of social ties and similarities among a population of students and the emotional response to content selection in a social network, in particular, surprise. The first study, which proposes an alternative method of investigating network bridges and focuses on novelty perception from receivers, supports the hypothesis that surprise is a proxy of novelty and, thus, bridging factors, i.e., tie strength and structural holes, can be predictors of

surprise response. The second empirical study builds on the finding that content selection in a social network environment is more dependent of receiver's emotional reaction than from factors associated with sender’s popularity or to a strong friendship bond between sender and receiver. The last study examines the optimal cognitive distance, between sender and receiver, measured by personal attributes that jointly with bridging factors predicts receiver's surprise response. The findings show that the performance of social network-based recommender systems can be improved by the delivering of novel and surprising recommendations based on prediction and not on randomness, which avoids the Social Echo Chamber Effect. This dissertation draws attention to the fact that the social data can be used to increase the cognitive distances among users, in order to deal with a set of new threats (e.g., at level of democracy/tolerance, conformity, cognition, “fluffy” innovation) that has been imposed by some web algorithms.

Keywords: Recommendation, Personalization, Structural holes, Ties strength, Surprise, Novelty, Homophily, Centrality.

CHAPTER 1

INTRODUCTION

Objectives

The Internet is a critical medium which gives us the opportunity to connect with all kinds of different people and provides access to information from all over the world. The almost instant access to global contents makes the Internet one of the chief powerful allies of globalization. Simultaneously, it acts as a glocal[1] medium. It means that its technologies transform the "global" into other shapes that meet the needs of local consumers. This double role seems to be built on how people trust its technology and in results obtained through Internet queries. In fact, it has become a general practice for a person to look for a particular solution over the Internet and then getting satisfied with the solution. They often believe the result they get is the best available for them.

This current state of things prompts two questions. a) What is the best result for an end-user?

Given the evolution achieved by some Web-based services (e.g., searching and recommendation), the answer seems to be related with personalized deliveries, which are intimately related with the increased performance levels of these services. In fact, with the growth and strengthening of the social Web and associated services and technology, users become treated as a selected audience by the content providers (e.g., media), which imposes, externally, a pre-constructed and imposed view. This view, which describes a current trend on Web, is "tuned" by the information obtained on the users' habits and interests.

In this vein, scholars have been addressing the advantages and disadvantages of personalization in several contexts, such as on media and by means of Web-based services. A common and generally accepted conclusion regarding disadvantages is that online personalization may isolate people from a diversity of viewpoints or content

[1] The term “glocalization” that describes a new outcome of local conditions toward global pressures, can be connoted with a successive development and challenge to the top-down hegemony implicit in the term “globalization” (Maynard, 2003).

(Nguyen et al., 2014). This fact emphasizes the meaning of living inside echo chambers, as argued by Sunstein (2009) in the book “Republic.com 2.0”.

With regards to its advantages, marketers will say that the value of personalization is in how to treat each person individually, with targeted content and offers that can appeal to their implicit or explicit needs. This is a scenario that makes even more acceptable the idea that almost everything is available over the Internet and translated into data[2] aimed to satisfy our needs in a more and more customized way. With this in mind, scholars and entrepreneurs started to look more attentively to the data collected in online social networks. Therefore, given the above question, the best result for the end-user would be related to more personalized results, but this means severe consequences to users in terms of diversity.

Moreover, there seems to be a gap which has not yet been addressed in the literature, where the benefit of using social data in personalization services is discussed. Next, I introduce this issue, by formulating the second question.

b) How different would the users' satisfaction level be if they could have access to the amount of information not shown due to personalization methods?

Some authors (e.g., Vargas et al., 2014) argue that the users' satisfaction level can be enhanced by means of the results' diversity. Agreeing with this viewpoint, the investigation undertaken in this dissertation shows how to achieve such goal. Chapter three presents some of the different personalization methods currently used.

Given the highly significant gain in popularity, online social networks became an important resource for recommendation (e.g., Ma et al., 2011; Bobadilla et al., 2013) and search (e.g., Mislove et al., 2006; Golbeck, & Wasser, 2007; Carmel et al., 2009). In particular, explicit user interactions have created an ideal test-bed for personalization. It was assumed that closely related people had similar interests, from which a

[2] The amount of data gathered globally has grown exponentially (McKinsey Global Institute 2011), as well as the value of the data produced in big social media platforms, such as Facebook, Twitter, LinkedIn, Pinterest, or even Foursquare. For example, Facebook processes around 500 terabytes of data every day, and their users exchange over 2.5 billion posts and upload around 300 million photos daily (Batorski, 2012). On the other hand, big data (unstructured information) have been marketed as one of the newest and promising business derived from the Internet. By mid-2008, Google already had in excess of a trillion unique web addresses indexed, while the number of queries entered into the search engine was around 2 billion every day. Thus, the data collection from online user behavior and status from online social networks, along with the development of the “Internet of Things” and the growing use of various sensors and devices connected to the Internet makes data even more special.

representation of the user and their social links in the network in relation with other users could be created (Boyd, 2008).

In this vein, scholars have been focused on the relevance of a recommendation based on social-influence relatively to similarities in past activities. It may mean we overcome the idea that recommender systems are "computerized oracles or black boxes" that give advice but that cannot be questioned (Groh & Ehmig, 2007, p.7).

However, the proposed approach to personalization also presents some constraints to receivers of information. I argue that the endogenous properties associated to people's behavior in their personal networks, such as the one characterized by homophily, may be extended into the social data used to improve personalization. Homophily represents the outcome~~s~~ of social processes which show that people of the same or similar groups tend to adopt similar behaviors and diverse behaviors if they do not share this common background. This social behavior strongly affects the creation and maintenance of ties with other people and the sharing of new information inside these groups of similar people. On the other hand, it is known that people generally seek out information and interaction that reinforces their private positions, and so, by avoiding engagement with difference, people become a source of their own "Echo Chamber" of information and viewpoints. This natural behavior of people in offline social networks is not seen as a threat from the perspective of the access to novel information. People in offline interactions have the freewill to reach different and socially distant individuals during their daily connections. This fact assures the access to diversity and novel information (Granovetter, 1973). However, the described scenario of interactions may change a lot in the online environment, notably, when the access to content is ruled by personalization based in social data. In such circumstances individuals get stuck in echo chambers without having the same "natural" liberty to access novel information. In general, people accept the results offered and trust the Internet.

Consequently, given the current trend of content personalization through the main forms of user interaction in the web – browsing, querying, recommendation – the flow of information, when grounded in data from similar people based on Web usage patterns, may satisfy the users' need for information, but often does not contribute to the diversity of their viewpoints (Golder & Yardi, 2010), which may reduce the quality of the service provided. Furthermore, it generates a low level of novelty in information access, which may constraint the enrichment of the meaning construction (i.e., for information interpretation). It means that, in spite of the enormous variety of sources of information, views and contents appearing online, people do not extract the whole benefit of its diversity. People are stuck inside their information bubble. Additionally, it is important to notice what is behind the data, either the social interaction that gives meaning to data, as the emotions' role in such interactions when there is information sharing.

Hence, in this dissertation, I examine whether or not the current use of social data may solve the problems related to the familiarity of contents accessed. It may even increase the difficulties. This problem is approached in this dissertation through the broadened concept of Social Echo Chamber Effect.

The introduction of the term "social" in the concept of "echo chamber effect", aims to explain this concept from a perspective that affects the final result of personalization. Accordingly, the use of data based on these kinds of attributes, rather than leverage innovation, may reduce the quality of the expected service. Therefore, this motivates the following research question:

How to use social data and avoid the Social Echo Chamber Effect?

In order to study complementary solutions and still benefit from the richness of social data, it is important to discuss the role of social data at a cognitive level. The field of Social Network Analysis (SNA) bestows a rich framework for studying such a problem. This is supported by rich theoretical and methodological contributions explaining the origin and consequences of such social dynamics, which also explain the Social Echo Chamber Effect and what solutions can be explored to counteract its effect. Moreover, the empirical knowledge provided allows the understanding of advantages and drawbacks of the use of social data, and what kind of social data should be considered.

In this dissertation I test the value of the information flow determined by individuals who are socially distant and have no redundant connections between them. This means, being connected by a bridging tie. Some of the advantages related to bridging ties (Granovetter, 1973; Burt, 1982) in the context of this dissertation, deal with the delivery of novel information, which may contribute to solving the Social Echo Chamber Effect.

However, despite the fact that researchers have been demonstrating the evidence of the delivery of novelty through the two main known bridging factors (i.e., weak ties and non-redundant structural holes), scholars have not been considering the receivers' side of this network mechanism. Thus, regardless of the rise of interest and empirical work on novelty related topics, as well as on the use of social data, there is a lack of research on the effects of information on receivers.

In this sense, it is important to develop a common methodological and conceptual base to define the emotional response to information access and related social theory with the bridge mechanism in social networks. This approach underlines the importance of understanding the interactions among human and network factors (e.g., emotional response, psychological characteristics, personal attributes and network structural conditions), and how they impact Web applications that use such social data. This dissertation attempts to put these three fields of study together: social network analysis, social psychology and information filtering.

In so doing, this work presents three empirical studies showing the relevance of network bridges as central nodes in defining the flow of novel information, and the importance of the emotional response in explaining receivers' options (i.e., content selection) and the perception of novelty. This has led to the formulation of the following hypotheses:

1) First empirical study: H1: *Surprise is a proxy of novelty*; H2: *surprise is elicited either when the information is delivered by one single bridging factor or by the composition of both.*

2) Second empirical study: H1: *there is a relationship between sender's popularity and content selection;* H2: *surprise response is associated with content selection;* H3: *surprise response is associated with the quantity of published content by the sender;* H4: *tie strength is associated to content selection, independently of whether the tie between sender and receiver is a bridge or not.*

3) Third empirical study: H1: *Surprise is elicited when sender and receiver share dimensions of status and attitude homophily and have similar interests in music and political views;* H2: *Surprise is elicited when sender and receiver are dissimilar;*

H3: *Surprise is elicited when sender and receiver are bridged by a weak tie, share dimensions of status and attitude homophily and have similar interests in music and political views;* H4: *Surprise is elicited when sender and receiver bridged by a weak tie are dissimilar;* H5: *Surprise is elicited when sender and receiver are bridged by non-redundant structural holes, share dimensions of status and attitude homophily and have similar interests in music and political views;* H6: *Surprise is elicited when sender and receiver bridged by non-redundant structural holes are dissimilar.*

Given these hypotheses, this dissertation has the following general objectives:

1) Find a data source associated to a solution that confirms the perception of novelty, in order to counteract the Social Echo Chamber Effect.

2) Analyze users' options in content selection. It means, knowing whether a sender's position as structural bridge is more relevant for content selection than a centrality position. It is also relevant to know the importance of the strength of the tie between sender and receiver for content selection. In sum, I want to know whether the end-users' behavior are based in the same assumptions as the ones applied by Web-based systems.

3) Identify personal and network dimensions to quantify the distance between senders and receivers, based on their similarities and dissimilarities, in order to provide two kinds of outputs: a) A methodology to identify new dimensions; b) Information to design predictive algorithms on surprise response.

And the following specific objectives:

1) Introduce an alternative method to study the perception of novelty given a delivery of information through a network bridge.

2) Analyze the influence of network dimensions (i.e., network centrality, structural holes, and tie strength) in individual's choices of contents.

3) Test a range of personal attributes combined with bridging factors (i.e., weak ties and non-redundant structural holes), to identify the optimal cognitive distance[3]

[3] As detailed in Chapter Eight, this concept was operationalized by Nooteboom (1992; 2005) stating the importance of differences in cognition between individuals in the context of novelty.

between sender and receiver of information, which is associated to the perception of novelty.

Reasoning in terms of networks and the method of network analysis have gained ground in many disciplines, such as social psychology, anthropology, or communications, to name but a few, I see that network model encourages scholars to use new cause/effect variables in their analysis. Some of them can be found through the properties expressed on communication networks, i.e., connectedness, integration, diversity, and openness (Rogers and Kincaid, 1981). I extend this view using the receivers' emotional response to identify their perception of novelty, when the access to contents is established by means of a network bridge. This study focuses on the analysis of relationships between people, but also in the characteristics of people, as well as on the established communication network.

This is the combination of topics that I found more adequate to investigate and introduce the concept of Social Echo Chamber effect in the context of social dimensions and dynamics in personalized recommendation. Moreover, it is demonstrated its impact in the quality of online recommendations. This dissertation test alternative social dimensions able to substitute the current flow of social data and so how to avoid the social echo chamber effect.

Moreover, the scholar's interest on understanding the reasons why communication networks emerge and the effects of communication networks seems also to have been growing, as stated by Monge & Contractor (2003). Regarding to communication networks, Rogers (1986) characterizes them as consisting of interconnected individuals who are linked by patterned communication flows.

Theoretical Background and Rationale

Scholars conceive that communication network analysis and structural analysis can be seen as intertwined, given the sharing of intellectual lineages though they have followed different paths of development and debate. Structural concepts, notably, have been introduced in diverse disciplines (e.g., linguistic, anthropology, sociology), since the beginning of last century (e.g., Saussure (1916/1966) within linguistic studies). It is in this context that Monge & Eisenberg (1987) debate with great detail three traditions, i.e., positional, relational, and cultural, which include most of the structural analysis of organizations and communication. The *positional* tradition departs from the idea that "positions and roles determine who communicates with who, and, consequently, the communication structure of the organization" (Monge & Contractor, 2003, p. 39). This 'static' view disregards individuals' activity in creating and shaping the organizational structure, as well as the role of their individual characteristics. It is considered that the organizational structure is set over a pattern of relations among positions. The assumption is that people occupying a given position are necessarily associated to behaviors, relations and sets of organizational roles. Although this tradition has its roots in classic works like Weber's (1947), "The theory of social and economic organization", or Homans (1958), "Social behavior as exchange", more contemporary works, like White et al. (1976) and notably Burt (1982), also have theorized about similar assumptions by developing the rubric of structural equivalence. One of the criticisms against this positional tradition is its inability to frame the way individuals take part in the creation and shaping of organizational structures.

The *relational* tradition is concerned with the communication linkages that are kept by direct communication. Monge & Eisenberg (1987) argue that this tradition is rooted in systems theory (e.g., Watzlavick et al., 1967), where the "denotation of the interconnections among systems components and the arrangement of the components into subsystems and supersystems" (Miller, 2011, p. 73) represents one of its hallmarks. In these systems the "mapping" of relationships among such components, when they are people and social groups, gain crucial relevance. Given this, Monge & Eisenberg (1987) emphasize the difference between positional and relational tradition, positing that a formal chart does not identify the actual systems of communicative relationships. The former refers to a prescribed flow of communication within an organization, given the formal organizational chart, while the latter mirrors the actual communication relationships emerging from the organizational system activity (Miller, 2011).

Finally, the *cultural* tradition examines symbols, meanings, and interpretations of messages transmitted through communication networks, highlighting the implicit, tacit and deeper meanings, as well as the shared values, in an organization. This tradition sees how meanings emerge from interaction and may constrain subsequent interactions (Monge & Contractor, 2003). It means that a common underlying structure determines individuals' interaction in organizations, going beyond a structural and individual view (Waldstrøm, 2001).

However, as Monge & Contractor (2003) claim, the three above-mentioned traditions cannot be seen in isolation, given that other theoretical mechanisms, like self-interest, contagion, and exchange, need to be considered. These mechanisms are particularly relevant to describing how people deal with linkage (i.e., creating, maintaining and breaking links) and so are pertinent for social networks formation. Furthermore, the wide range of social network theories is often related to the topics of user incentives (e.g., friendship, appreciation), but also with the theories of self-interest that debate on people's choices driven by preferences or desires given what they believe to be an acquiring of personal benefit (Monge & Contractor, 2003).

Furthermore, benefits acquired by the network interaction are not often thought from the cognitive and emotional viewpoint. Psychological attributes together with structural factors have not been deeply analyzed in this context. The arguments have been mainly focused on the gains explained by the theory of self-interest, comprising other theories like *social capital,* which broadly discusses the potential benefits retrieved from communication networks in which people are key actors. In this context, bridging factors attributed to structural holes (Burt, 1992) are used as a mechanism that gives access to such personal profits. Nonetheless, rather than considering bridging factors only as a hinge that gives access to a spectrum of benefits through the agent of the transaction (usually taken as the beneficiary), they can also be analyzed from the perspective of receivers' benefits. This latter perspective involves two important reflections. Firstly, the information contained in the delivered content that stimulates the construction of meaning on the receiver may act as a proxy of the psychological characteristics of the sender (e.g., personal traits). Secondly, the benefits for receivers may be due to the perception of novelty and also by the surprise elicited. This motivates new research on the role of individual attributes and psychological characteristics in the flow of information in social networks.

On the other hand, scholars often refer to "spread" (e.g., Bakshy & Rosenn, 2012) to denote flow (e.g., of information) or movement in a social network, whereas the sender influences the receiver (also known as adopter, in diffusion literature). This influence is often attributed to the strength of the tie, or homophily-related effects. This dissertation does not address the study of diffusion, but analyzes how the elicited emotions (i.e., surprise) are intertwined in the interaction of social networks given the structural and relational properties and individual characteristics.

Additionally, it is known that the Internet is a fast and ubiquitous channel of communication, but it is also important to understand how people are connecting and what they are saying (Watts, 2007). In particular, social networks offer an open window to observe people's behavior, their tastes, moods, health, and the impact of person's structural position (in a network) over these dimensions (Lewis et al., 2008). Moreover, with the emergence of computational social science it is becoming feasible to collect and analyze massive quantities of data. However, it seems that the leverage of new opportunities to study human behavior is more related to the value of interdisciplinary fields, than with the storage of massive data describing minute-by-minute interactions and locations of entire populations of individuals (Lazer et al., 2009). In fact, nascent interdisciplinary fields and questions are now appearing from computational social science, as well as from other fields such as neuroscience and social psychology, which highlights the need and opportunity for more crossing-disciplinary studies.

It is within this logic that the link between the research questions set on the scope of SNA and the findings of the empirical work of this dissertation are built.

Analytical approach

This dissertation debates the broadened concept of Social Echo Chamber Effect to deal with the cognitive factors that are intimately associated with personalization constraints. These cognitive factors are then related to the use of network data. In order to study this problematic and find an alternative solution, the empirical work presented in this dissertation tests the relationship between emotional reaction (i.e., surprise) and several network dimensions. The goal is to find the adequate source of data that counteracts the effect of social echo chamber.

Although this work applies SNA theories and methodologies to study the problems outlined in the context of personalization, this dissertation also discusses other findings achieved within the SNA field.

The social networks approach offers theory and methodology with applications to all levels of observation of the network actors (Marsden, 1990). This perspective has favorable analytical properties to measure how individual choices may be affected by factors related to individual attributes and relational properties in an inherently structural framework. Nevertheless, individual choices also are affected by emotional reaction. The examination of receivers' emotional reaction introduces a weighting measure among variables represented by structural (e.g., structural holes), relational (e.g., tie strength) and endogenous properties (e.g., homophily), which highlights the pertinence of the relationship between network properties and emotions elicited.

Taking this approach, it becomes possible to study the use of social data by means of novelty perception to find an answer to avoid the Social Echo Chamber Effect, as proposed by this dissertation.

Summary and Preview of Chapters

This chapter provided a conceptual overview of the problem debated in this dissertation and briefly defined the constructs of interest of each empirical study undertaken.

In Chapter Two, I provide a review of extant literature related to some cognitive factors in the online environment. This dissertation debates the use of an emotional reaction, i.e., surprise, related to a cognitive effect, i.e., novelty, to propose a solution for digital media systems. Thus, it is relevant to presents an overview of the literature on cognitive factors in the scope of online human-computer behavior. A particular attention is given to the process of construction of meaning due to its relationship with subconscious activity stimulated by the emotion of surprise. Equally, it is relevant for the argument of this dissertation, the association between novelty and surprise.

In Chapter Three, I present a review of existing literature related to Web-based services, in order to introduce trade-off between the evolution of some of these Web-based services and cognitive factors, and like that, contextualize some failures or abandonment of some Web technologies. In Chapter Four, I present and explain the concept of Social Echo Chamber Effect, introduced in this dissertation. In Chapter Five, I present a review of extant literature and main social network variables used discussed in the empirical studies of this dissertation.

In Chapter Six, Seven and Eight, I present these three empirical studies, which address different aspects of the research goals discussed above. This chapters are designed almost as stand-alone articles, meaning that each is written with introductory material; a description of the measures, data, and analysis; a presentation of the results; and a discussion of that specific study's findings. The previous chapters are also aimed to introduce the background of these studies.

Finally, in Chapter Nine, I present an overall conclusion which seeks to synthesize the main findings across the three empirical papers and articulate some general considerations for assessing the project as a unified whole.

CHAPTER 2

COGNITIVE FACTORS IN THE ONLINE ENVIRONMENT

2.1 Overview

Most traditional Web-based applications have focused on improving the productivity or decision making of the individual user through personalization. The emphasis has been on providing the tools and data necessary to fulfill a specific job function, such as searching, browsing or recommending. Emotions are also considered, such as in searching, given the relevance of avoiding users' feelings of regret or frustration. Meanwhile, other cognitive factors, i.e., emotional reaction, associated with social network dynamics, also play a relevant role in user productivity and in the interpretation of information.

This chapter presents an overview of the literature on four cognitive factors (i.e., Intellectual styles, Construction of meaning, Unconscious role of cognition and emotions, Emotions and novelty), which are interconnected and intimately related to how the online environment and its objects may interact differently among different users. This is important to interpret the online human-computer behavior and to uncover possible constraints hidden behind such interactions. The last section of the chapter overviews the concept of emotion in general and the relationship between surprise and novelty in particular. This is particularly important in the context of this dissertation because it justifies the method applied to study the mechanism of bridging from the receiver’s viewpoint. This method is based on the use of surprise as a proxy of novelty.

This chapter is organized into four sections. The first is named *Flow and Intellectual styles in online human–computer interaction* and starts debating how cognitive thinking style influences users’ behaviour. The second is called *Construction of meaning* and introduces the concepts of meaning and meaning construction. It draws our attention to the process of information interpretation, and the association between meaning and emotions and how meaning emerges from context. Next, in the section of *Unconscious role of cognition and emotions*, the role of the unconscious in cognition and emotions is debated, e.g., primary emotions – like surprise, which are typically associated with unconscious processes – and the association between emotions and specific cerebral hemispheres. Finally, in *Emotions and emotional perception of surprise and novelty*, an overview of emotions, emotional perception and novelty is presented in order to explain the association between novelty and surprise. The aim here is to justify the use of surprise as a proxy of novelty in the three empirical works of this dissertation.

2.2 Flow and Intelectual styles on online human–computer interaction

Flow theory has its roots in psychology and is used to address optimal user experiences with personal computers (e.g., Ghani, 1995) and the World Wide Web (e.g., Chen, 2000, Novak et al., 2000). As a construct for describing more general human–computer interactions in online environments (Trevino & Webster, 1992; Trevino & Ryan, 1992), flow was important for understanding consumer use of the Web (Hoffman & Novak, 2009). Flow can be defined as “the state occurring during network navigation which is: (a) characterized by a seamless sequence of responses facilitated by machine interactivity, (b) intrinsically enjoyable, (c) accompanied by a loss of self-consciousness, and (d) self-reinforcing” (Hoffman & Novak, 1996, p. X). Given this state of mind, the user forgets everything else around him, like time (Novak et al, 2000). Thus, flow represents a state of consciousness where a person is so absorbed in an activity that s/he excels in performance without consciously being aware of his or her every movement. The use of this theory has been applied as a way to understanding human behavior with computers and thus inform better ICT[4] design, training and use (for a review see Finneran & Zhang, 2005).

Novak et al. (2000, 2003) state that there is more evidence of flow for task-oriented activities than for experiential activities, but that there are flow experiences in both types of activity. Furthermore, online customer experiences are positively correlated to “fun, recreational and experiential uses of the Web”, and negatively correlated to work-oriented activities. This definition and the existence of such flow experience in the Web environment, was empirically tested by Chen (2000), who contends that Web activities provide enjoyable experiences to Web users improving the quality of their psychological well-being. The flow in the Web environment is presented in this context as being related to functional categories, i.e., researching on the Web,

[4] Information and communications technology.

information retrieval, participating in discussion groups, e-mailing, creating Web pages, playing games, and chatting.

However, as stated by Hoffman & Novak (2009) “the consumer Internet has evolved from a few directories and online storefronts into a vast, sophisticated network of information stores that millions of people interact with on a regular basis.” Even more nowadays with the maturity of the “Web 2.0” and the impact of social networks on users’ habits and flow of information. As matter of fact, ‘‘members in virtual communities differ from general Internet users in that virtual community members are brought together by shared interests, goals, needs, or practices’’ (Chiu et al., 2006, p. 1875).

Flow has been examined as antecedents of behavioral intentions and behaviors, such as related to the influence of flow on continued use of mobile instant messaging (Zhou & Lu, 2011), the impact of instant messaging flow experience on exploratory behavior (Zaman et al., 2010), the importance of flow experience as a mediator that produces indirect effects in predicting the social network sites games continuance in the model (Chang, 2013), or the contribution of both knowledge seeking and knowledge (contributing in the context of Web 2.0 virtual communities) to flow, and also to employees’ creativity (Yan et al., 2013).

In this vein of investigation, Vinitzky & Mazursky (2011) argue that beyond the effects of online human-computer interaction (e.g., Novak et al., 2000), it is important to consider users’ personal differences in their cognitive thinking styles and that cognitive thinking styles influence users’ behavior. The results presented by these authors show that intuitive thinking style promotes associative thinking and pleasure, thus, the more pronounced this style is, the higher is users’ perception of interactivity of a Website. In turn, systematic thinking style does not promote exploratory behavior or the perception of interactivity. Additionally, differences in cognitive styles influence the amount of information sought to support the decision-making process and the corresponding number of alternatives to be considered by the individual (Hunt et al., 1989; Driver et al., 1990).

Cognitive styles refer to consistent individual differences in how individuals perceive, think, solve problems, learn, take decisions and relate to others (Witkin et al., 1977). These psychological dimensions represent consistencies in how individuals acquire, evaluate, organize and process information, and guide their performance in information processing and creative tasks, through relatively stable mental structures or processes (Messick, 1984; Myers & McCaulley, 1985; Aggarwal, 2013). Thus, cognitive styles are understood as an internal preference of the individual for using a unique type of thinking (Sternberg, 1998), whose pattern tends to be stable over time and in different situations and is independent of the level of intelligence (Perkins, 1981).

More recently, the concept of intellectual styles has been seen in scientific literature as an umbrella term that covers closely related constructs such as "cognitive styles," "learning styles," "teaching styles," and "thinking styles". One example of this can be found in the work of Zhang & Sternberg (2009). Such terminology basically intends to explain why different people succeed in different professional and organizational settings. In this regard, it was thought for a long time that innate abilities justified differences between high-achievers and less successful peers. However, research has shown that individuals have different intellectual styles that fit in varying types of contexts and problems (Furnham, 2011). Thus, despite the fact that literature uses different terminology to explain “style”, it is accepted that intellectual style “refers to one’s preferred way of processing information and dealing with tasks. To varying degrees, an intellectual style is cognitive, affective, physiological, psychological and sociological” (Zhang & Sternberg, 2005, p. 2). Indeed, apart from some confusing literature on all sorts of styles (Furnham, 2011) most scholars believe that styles are primarily a function of ability and personality (Zhang & Sternberg, 2000, 2005).

The questions of cognitive styles is of significant importance, both scientifically and practically (Zhang and Sternberg, 2009), but despite the growing interest in this field of study, it is still a relatively neglected concept in several areas, like business and management fields (Amstrong et al., 2012), or Web-based systems (e.g., Kao et al., 2009, Ocepek et al., 2013).

Various studies demonstrate the significance of compatibility between styles and task or activity characteristics (Epstein, 1994, 2003; Hogarth, 2002; Kahneman, 2003; Novak & Hoffman, 2009). Considering the business and management fields, Amstrong et al. (2012) present an extensive literature revision, where they conclude that cognitive style can be a critically important indicator of vocational orientations, vocational choice, job selection, job level and work performance. Furthermore, cognitive styles are likely to have an impact on aspects of perception and communication in teams, membership formation, group norms and deviancy, individual versus group goals, team leadership, group problem solving and decision making, and group conflict.

In the context of We-based systems, it is argued that human-centric recommender systems are an adequate solution to satisfy users' new needs that are more and more specific and based on products located in long tails (Gretzel et al., 2012). The authors posit that the success of a specific destination recommender system depends on the ability to anticipate and respond creatively to transformations in the personal and situational needs of the users.

Cook (2005) argues that one way of maximizing learning in web-based environments is to adapt web-based environments to suit specific cognitive styles. For example, individuals with analytical styles in environments with no clear structure, which are somewhat informationally disoriented as well as socially isolating, are more able to benefit from their own structure. They require less external motivation and social support, which can be used as an advantage over individuals with holistic styles. Conversely, holistic styles are at an advantage when the environment is characterized by settings with explicit guidance and structure, external motivation and social interaction (Chen and Macredie, 2002). This vision of an adaptive learning environment is based on the idea of "one teacher for one student" (Woolf, 2009). This is a statement based on the constructivist learning theory, which supports the idea that knowledge is constructed by the student individually through his interactions with the learning environment (Rovai, 2004). Students can select their own material and learning resources by themselves, according to their preferences. However, this process may cause a cognitive overload or stress on students. The overload may originate from paying too much attention to selecting the appropriate presentation of learning topics (Mayer & Moreno, 2003). Stress can be caused by inappropriate multimedia material selection (Chen & Sun, 2012). In order to avoid this, a recommender system may recommend the appropriate learning materials taking into account student's preferences while guiding them through the learning process (Vogel-Walcutt et al., 2011). Accordingly, Ocepek et al. (2013) propose an adaptive constructivist learning environment that recommends learning objects. The goal is to relate the combination of different learning style models with the preferred types of multimedia materials in order to select appropriate multimedia types for particular students. The results show that the learning style model of hemispheric dominance is the most important criterion in deciding if students prefer different learning multimedia materials. It was also found that the most of students still learn using textbooks and books.

Kao et al. (2009) argues that human factors, such as thinking style (an affective factor) should be incorporated in the design of search engines, because it influences search target settings and search behavior. Additionally, it can be used with or without data mining techniques to identify user search patterns for predicting search intentions. The relevance of this suggestion seems to rely on the fact that search results are sorted using relevance-ranking mechanisms, which do not provide significant or structured presentations in a friendly way to help users quickly comprehend the retrieved information (Kao et al., 2009).

Other approaches state that team composition based on members' cognitive styles explained differences in performance between teams. It influences both the strategic focus that a team forms, as well as strategic consensus. Diversity is categorized here in terms of race, ethnicity and gender (Aggarwal, 2013). Other studies examined the effect of interpersonal differences in thinking style on online consumer experience (online purchase process) (Vinitzky and Mazursky, 2011). The findings indicated that systematic cognitive thinking style is correlated to search motivation. It means that online stores with an environmental distracter may be less accurate at capturing purchase attention from people with this kind of cognitive thinking style. In this study, the authors differentiated the cognitive thinking style between systematic cognitive thinking style and intuitive cognitive thinking style[5]. They emphasize the need to consider consumers' shopping environment and personal differences in their cognitive thinking styles. Thus, the rise in the consumers' satisfaction level and their loyalty to the site seems to be related to the structure of the site, its contents, and its advertising information with regard to consumers' shopping environment.

Zhang & Sternberg (2005) classified all major style constructs in styles literature in three types: Type I is associated with right-hemispheric styles being indicative of higher levels of cognitive complexity. Type II is associated with left-hemispheric styles and denotes lower levels of cognitive complexity. Finally, Type III, which manifests the

[5] The authors report that systematic thinking is related to a person's tendency to analyse information and reality in a rational, consistent, and multilevel way. Intuitive thinking is related to the individual's tendency to organize information globally and to make decisions after he/she has already formed, developed, and understood the entire context of the required decision.

characteristics of either Type I or Type II styles, depending on the stylistic demands of a specific task (Furnham, 2011).

Cerebral predominance – be it right or left – is then intimately related to style thinking. Left cerebral hemisphere is intimately related to analytic, rational, and sequential information processing, and the right cerebral hemisphere specializes primarily in intuitive and simultaneous information processing (Armstrong, 1999).

Fecteau et al. (2004, p. 551), say in this context that "word reading is one of the most strongly lateralized, showing a left hemisphere advantage" and that the left hemisphere displays some advantages such as helping in tasks that involve word reading and that are related to visual stimuli. This is to say semiotic activities, or use of explicit information. Hence, we may see the left hemisphere as the basis of a linguistic frame, being the language a semiotic tool applied namely to construction of meaning and meaning exchange in imagined or real social interaction (Holtgraves et al., 2007).

Yet, though certain cognitive activities are intimately related to a certain hemisphere, this does not mean that the other is not able to actively participate in the interpretation of information. Both sides of the brain participate simultaneously in the construction of meaning, albeit with different weightings of activity (Fecteau et al., 2004). These authors argue that "the right hemisphere shows as much evidence of reading words unconsciously as the left hemisphere. Thus the classic left hemisphere advantage in word reading is likely only to be an advantage of conscious access to words presented to that hemisphere" (p. 562).

As matter of fact, it has been reported that the process of conscious thinking is related to explicit information, which is typically associated with activities such as word reading (Fecteau et al., 2004). When reading, the unconscious activity of the individual makes use of the implicit information to achieve meaning, which justifies the fact that what is explicit through words does not mean the same to everyone. As a result, the response[6] related to information access plays a relevant role in how people elaborate

[6] Neuman (2004) defines meaning as "the systems specific response to a signal", and meaning-making as "the process that yields the systems specific response to an indeterminate signal". In this regard, the author clarifies that 'response' in this context is not associated with the sense of 'stimuli–response', which could be wrongly related to 'behaviorism'. It intends to describe an *interaction* with the environment. Behaviorism was established with the publication of Watson's classic paper "Psychology as the Behaviorist Views It" (1913). This theory of learning is based upon the idea that all behaviors are acquired through conditioning. See more here: http://psychclassics.yorku.ca/Watson/intro.htm.

new perspectives and viewpoints. This, given the level of interpretation achieved and associated richness of the construction of meaning.

On the other hand, unconscious thinking is claimed to be related to implicit information (Ekstrom, 2004), which is triggered by certain types of stimulus (e.g., specific emotions) (Scarantino, 2005)[7].

In this view, users' online experiences are strongly dependent on users' characteristics, which have been explored by researchers and entrepreneurs to improve the human-computer interaction. The lessons learned show that flow is less dependent on user-machine interaction and is also influenced by interactivity-community/ recommendation interplay. This seems to be in line with the current trend of personalization. Nonetheless, the forecast improvement in quality of web experience for users through online personalization seems disappointing from the consumer's viewpoint (Lee et al. 2009). People seem to feel their freedom is threatened when they are given these kinds of recommendations.

2.3 Construction of meaning

The concept of information has several senses (Collier, 1990), but is often associated with Shannon's (1948) statistical definition of information, which separates information from meaning. Because of this, the concept of information has been frequently seen from the perspective of its quantification (Aczel & Daroczy, 1975; Cover & Thomas, 1991). In this sense, the debate ranged from the quantification of the information included in a piece of data to the measurement of the information yielded by one event (Cover & Thomas, 1991). Though, considering that a bit of information is like ''a difference that makes a difference'' (Bateson, 1972, p. 315), it is correct to analyze the interdependence between information and meaning, i.e., they are closely related (Neuman, 2006). However, it would be misleading to consider that the meaning of a message can be reduced to information content, just because meaning is also about the information carried in the detected message. Here, the meaning-making that acts as a procedure for extracting the information conveyed by a message (Neuman, 2006), may not be able to extract the meaning from the context, which could complete the meaning of the message. As Neuman (2006) underlines, in order to understand how

[7] In view of brain areas assessment in this regard, is reported that no physical or chemical measurement of brain activity is a direct measure of meaning (Freeman, 2003).

meaning emerges from the context, first it is necessary to find better ways to model meaning-making, which is an important issue not only for people in information sciences, but also for the field of artificial intelligence.

From a human approach, sometimes, one word it is enough to maintain the dialogue between two people; however, a third person may have difficulties in catching the meaning of the message. The issue here is how to understand the meaning that emerges from the context that is comprehended by the two people, but not by the third person. As pointed out by Neuman (2006) "we still do not have a satisfactory answer to the question of how meaning emerges in context" (p. 1447). This is a relevant question, not only from a theoretical approach, but also for computation (e.g., in artificial intelligence research). As a result, different readers will get a different meaning from that, which is conveyed by the words, as well as other semiotic symbols. Each person acting as a receptor uses his own background and expectations when interpreting information (Freeman, 2003).

Cognitive scientists studying meaning have achieved many similar sorts of ideas as those studying vision. It is assumed that there is a considerable difference between the visual information transduced by the eyes and the information that the brain subsequently computes from it. Similarly, the information contained in linguistic input does not fully describe its emergent meaning. In this sense, words and varieties of linguistic structure have no intrinsic meaning; they are used by speakers to actively construct meaning (Coulson, 2006). This explains the complexity of detecting novelty, in particular in computation through linguistic events. Langacker (2000) refers to this problematic in the context of novel expressions in these terms:

> "when a novel expression is first used, it is understood with reference to the entire supportive context. The speaker relies on this context, being able to code explicitly only limited, even fragmentary portions of the conception he wishes to evoke. Usually, then, the expression's conventionally determined import at best approximates its actual contextual understanding. (…) It does not contain or convey the intended meaning, but merely furnishes the addressee with a basis for creating it." (p. 15)

Furthermore, Chandler (2005) contends that “meaning does not reside *in* a text but arises in its interpretation, and interpretation is shaped by socio-cultural contexts”. This idea agrees with the saying of Paul Valéry (1957, p. 1597) who states that “there is no true meaning of a text”. When a text is interpreted by its receiver, it is already free of the contextual support of the author (of the text) to be formalized in the cognitive contexts of the receiver. The message immersed in a narrative is then passed from the sender to the receiver in a continuum of contexts, both conscious and subconscious, and both converging in the embodiment of 'meaning' (Eco, 1990).

Meaning is equally associated to intensely positive experiences and then to the eliciting of emotions (King & Hicks, 2009; Keltner & Haidt, 2003). In this vein, a recent theory on the *Broaden and Build Model of Positive Emotions*, Fredrickson (1998, 2000) highlights the relevance of positive emotions for health and well-being. The author claims that positive emotions have a lasting undoing effect on negative emotions. Thus, strategies that cultivate positive emotions, like finding positive meaning, prevent or solve problems such as anxiety. In this regard, Schwarzer & Knoll (2003, p. 13) say that there is empirical evidence attesting “the fact that meaning and positive emotions help to restore an individual’s world view and may build additional personal resources”.

Meaning can also be seen from a social construction perspective through an information sharing environment. In this context, information sharing disseminates information that holds the same meaning to everyone (Miranda and Saunders 2003)[8]. Here, when a group member has equal access to information, it supports the social constructionist perspective that states that meaning is socially constructed during information sharing. However, this does not mean that agreement and shared meaning signifies the same. When agreement is reached, the offered meaning can become part of the common ground and then, the agreed-on interpretation of the situation is achieved (Bossche et al., 2006). This is the case of the process of building a shared conception, e.g., to solve a problem, which starts with the way the articulation of personal meaning is taken up in the social setting (Stahl, 2000). On the other hand, shared meaning refers to events wherein everyone holds the same meaning. Linguistic events where meaning is situational are an example of that (Ricoeur, 1981).

[8] The author bases his argument on Berger and Luckmann’s (1966) work on social institutions which proposes that institutions experienced as an objective reality are in actuality, social constructions, and on Schutz (1967), who emphasizes the cognitive processes underlying such social constructions.

It is known that the access to online content is characterized by a process of reasoning and interpreting. This process of interpreting is intimately related to meaning (Sommer et al., 1998), which can be detected or constructed. It is detected when it is related to pre-existing beliefs, and constructed, i.e., actively molded, when it is engaged in a constructive process to come to a sense of meaning (King & Hicks, 2009).

The detection of meaning can be observed in the way people perceive life and hypothesize about its meaning. Here, “meaning detection refers to those times when the data from the world are essentially (and perhaps quietly) saying ‘Yes’ to that hypothesis” (King & Hicks, 2009, p. 318). Nonetheless, the detection of meaning is not limited to a passive reception of meaning. It is a personal process that converts information into a personal perspective making it present when the event 'makes sense'. Because this is a personal process, it can involve new experiences that confront pre-existing assumptions, i.e., that fit with pre-existing beliefs and expectations (King & Hicks, 2006; Heine et al., 2006).

In contrast with meaning – related to detecting – meaning construction involves the cognitive action of searching for satisfactory answers requiring a revision in the meaning structures of the individuals (King & Hicks, 2009). It enables the interpretation of information, being dependent on the conscious and unconscious processes of thinking (Bargh, 2011). Thus, contrary to meaning detection, meaning construction is about awareness and intentionality involving an effortful process (King and Hicks, 2009) to avoid a threat in the individual sense of meaning (Heine et al., 2006). The motivation to keep this brain process, whose purpose is to maintain meaning, may correspond to the individual's awareness about the gap between experience and expectation (King & Hicks, 2009).

A daily example of construction of meaning can be found in the complex interplay between linguistic and nonlinguistic knowledge (Sperber & Wilson, 1995) that people do to define a week day. In this case there is a contextual dependence to understand the meaning of weekend, which is impossible to have without comprehending first the structure of the week and the respective cultural (and economic) knowledge associated with it (Fillmore, 1976).

2.4 Unconscious role on cognition and emotions

It was with Freud and Jung's work that the first attempts were made to systematize and understand the complementary relationship of the conscious and unconscious and their processes in cognition. Freud had the foresight to look to the brain for answers, but the mechanistic understanding available at the time was not enough to advance his studies. In this regard he mainly focused on unconscious *thoughts*. Jung, on the other hand, questioned the ability of the conscious to elaborate all the complex information from images and ideas (Ekstrom, 2004). With a gap of more than 50 years, today's cognitive scientists attempt to fully understand cognition in a holistic way. Nonetheless, it is relevant to note that old and new models connect in some obvious ways.

Bargh & Morsella (2008), claim that one reason for the lack of comprehension of this side of cognition was related to the meaning attributed to the term *unconscious*. The author says that "the earliest use of the term in the early 1800s referred to hypnotically induced behavior in which the hypnotized subject was not aware of the causes and reasons for his or her behavior" (p.3), as was reported by Goldsmith (1934). Also other scholars like Darwin (1859) in "On the Origin of Species", or Freud, applied the term unconscious to classify a non-intentional and deliberate selection (i.e., "unconscious selection") or to refer to behavior and ideation that was not consciously intended or caused (i.e., unintended behavior), respectively. In fact, the notion that certain universal truths came to stay is an idea strongly rooted in Western culture.

Given that, it is claimed that scholars in psychology, e.g., in the judgment and decision making (JDM) field, supported their cognitive research in the Cartesian tradition. It means that for these scholars, reasoning and judgment are an exclusively conscious activity, and that conscious short-cuts could equally be used under time pressure e.g., in heuristics (Lassiter et al., 2009).

It was in this scope that the cognitive linguists and philosophers, George Lakoff and Mark Johnson, challenged the premises of cognitive science, proposing a new approach to the understanding of the unconscious. They call it 'the cognitive unconscious', a concept which was developed in their book "Philosophy in the Flesh" (Lakoff & Johnson, 1999). The authors base their analyses on cited studies in neuroscience, cognitive linguistics, and neural modeling. They wrote on this subject:

> “Cognitive science is the scientific discipline that studies conceptual systems. It is a relatively new discipline, having been founded in the 1970s. Yet in a short time it has made startling discoveries. It has discovered, first of all, that most of our thought is unconscious, not in the Freudian sense of being repressed, but in the sense that it operates beneath the level of cognitive awareness, inaccessible to consciousness and operating too quickly to be focused on.” (p.9)

Exploring other approaches but still uncovering the role of the unconscious in cognition, scholars (e.g., Nisbett & Wilson’s, 1977; Ekstrom, 2004; Bargh, 2011) have been presenting the relevance and tangibility of unconscious thinking. For example, some studies show that while conscious thought is considered to be better for simple choices, unconscious thought does the deliberative work related to complex decision better (Bargh, 2011). In this regard, Dijksterhuis & Nordgren (2006) posit that the best strategy would be to consciously encode all of the relevant information and then let the unconscious do its task.

On the other hand, it seems that the conscious also uses already-existing unconscious motivational structures to pursue its goals (Dennett, 1995), illustrating the flexibility and adaptability of the unconscious processes (Bargh & Morsella, 2008). Furthermore, when conscious attention is diverted by a secondary task, this does not change the similarity on judgment outcomes produced by conscious and unconscious deliberation (Bargh, 2011).

These and other works revised in Bargh (2011) show that people think unconsciously as well as consciously, namely in the domains of judgment and decision making. Additionally, primary emotions (surprise, happiness, fear, anger, disgust, and sadness) (Izard, 1991) are not typically frequently activated in a consciously controlled real life. Emotion is an innate and unconscious process that has the ability to deal with cognitive processes (and problems) that do not require conscious attention (Scarantino, 2005). Centered mainly in a small set of sub-cortical brain systems, the emotion is like a biological sensor that alerts us to an opportunity e.g., danger, food, novelty, telling us to stop doing what we're currently doing, in order to attend to this challenge. In this sense, emotion has the tendency to respond strongly to high contrast information, and to just be vigilant in steady states or subtle changes. Further, different emotions are mainly related to specific hemispheres. For example, boldness is processed principally in the left hemisphere, and anxiousness in the right hemisphere (Siegel, 1999). Moreover, unconscious processing of emotional information is mainly subsumed by a right hemisphere sub-cortical route (Gainotti, 2012).

2.5 Emotions and emotional perception of surprise and novelty

Scholars agree that emotions encompass several aspects, such as information-processing components, response components, and regulatory components (Oatley & Jenkins, 1996), and are even directly related to the situational meaning. Different emotions occur in response to the meaning structures of given situations (Frijda, 1988; King & Hicks, 2009) that influence cognitive activity. They are dependent not only on the situation's characteristics, but also on the individual's perceptions (Frijda, 1988), i.e., the existence of an emotion itself depends on the perceiver (Lindquist et al., 2012).

There is a general consensus that basic emotions are psychologically primitive. Primitive means that they must originate in sub-cortical brain structures. In this sense, the neocortex may be involved in emotion processing, related to higher order structures. Thus, "basic emotion should be discrete, have a fixed set of neural and bodily expressed components, and a fixed feeling or motivational component that has been selected for through longstanding interactions with ecologically valid stimuli (e.g., the subjective feeling and motivational component of fear is what it is because this response has historically been most adaptive in coping with typical fear elicitors)." (Tracy & Randles, 2011, p. 398). This view is accepted by four prominent scholars in this field, Ekman & Cordaro (2011); Izard (2011); Levenson (2011); and Panksepp &Watt (2011).

In this regard, Izard et al. (1974, 1977) elaborated a scale of ten primary emotions, named "Differential Emotions Theory" (DES scale), which has been revised a number of times subsequently. The DES scale is composed by the emotions of: Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, and surprise.

In the revision undertaken by Tracy & Randles (2011) of the four above mentioned scholars, it is reported that the four authors agree in five common primary emotions. It includes a positive emotion named happiness (Ekman and Cordaro; Izard), enjoyment (Levenson), or play (Panksepp and Watt); and three distinct negative emotions: sadness (labelled grief by Panksepp and Watt), fear, and anger. Because there is some controversy on some emotions, such as surprise, contempt, and lust, they are out of the final list (Tracy & Randles, 2011). Although it is agreed that there is no sufficient evidence for their inclusion, there also is no reason to disregard them as basic emotions.

The emotion itself starts with a process of cognitive appraisal[9] that is related to the result of how people evaluate their continuous transactions from perceptual stimuli (Schmidt et al., 2010). As explained in Finkenauer et al. (1998), the appraisal process is influenced by several factors, such as antecedent personal characteristics – e.g., beliefs about oneself and the world – prior experiences and expectations, or even attitudes and self-concepts.

In this scope, Scherer's (1984, 1986) theory of components process model of emotion outlines a mechanism for the ongoing appraisal of environmental events. The author presents specific hypotheses regarding the pattern of meanings that will precede particular emotional states, as explained in Leventhal & Scherer (1987). Scherer proposes five types to check (or to evaluate) the emotional response, which includes a check for: 1) Novelty; 2) Intrinsic pleasantness; 3) relevance and /or conduciveness to meeting goals or plans; 4) ability to cope with the perceived event; 5) Compatibility of events (including actions) with self-concept and social norms. In this sense, it is suggested that surprise can be seen as a positive outcome of the novelty check – and so, a specific consequence of the appraisal of novelty (Finkenauer et al., 1998) – i.e., Novelty “determines whether there is a change in the pattern of external or internal stimulation, particularly whether a novel event occurred or is to be expected” (Leventhal & Scherer, 1987, p. 15). Another example is enjoyment, which is seen as a positive outcome of the intrinsic pleasantness check.

In short, emotions can be defined “as episodic, relatively short-term, biologically based patterns of perception, experience, physiology, action, and communication that occur in response to specific physical and social challenges and opportunities" (Keltner & Gross, 1999, p. 468). An emotion represents a complex array of psychophysical stimuli that arises spontaneously 3000 times faster than rational thought. It invokes either a positive or a negative response and typically a physical expression

[9] Appraisal is defined by the Merriam-Webster dictionary as an evaluation of worth, significance, status or estimate. Appraisal is achieved by monitoring and evaluating an event associated with emotional states (Smith & Ellsworth, 1987). The process of cognitive appraisal is influenced by personal characteristics, prior experiences and expectations, as well as behavior (Finkenauer et al., 1998).

(Tang et al., 2011). Thus, the emotional response (e.g., autonomic reactions) prepares the organism for action (Leventhal & Scherer, 1987). Cognition, on the other hand, corresponds to the appraisal or "evaluative perception" of the implications (positive or negative) of the stimulus for the organism. This appraisal operates at a simple sensory level or on the level of complex, conscious reasoning (Lazarus 1984), though, this is essentially a functional analysis, whereby it neglects the type of processing occurring.

In response to the question of which comes first, emotion or cognition, Leventhal & Scherer (1987) argue that it is difficult to conceive an emotional reaction totally unconnected from perceptual or cognitive reactions and that "emotion" and "cognition" "are always intertwined in emotional behavior and emotional experience" (Leventhal & Scherer, 1987, p.23).

Emotions in general, and surprise in particular, may be classified according to different grades of relevance and used in practical applications, notably by basing them in the correlation between novelty and surprise (Baldi & Itti, 2010).

Surprise is an important attractor of human attention (Itti & Baldi, 2009) and appears to be stimulated in situations in which an activated schema (Schuetzwohl, 1998) is interrupted by a novel, unexpected turn of events (Teigen & Keren, 2003). As shown in the empirical work of Reisenzein (2000) there is good evidence to trust that high surprise ratings are associated with low probabilities, and vice versa, though, as shown by Teigen & Keren (2003), the emotion of surprise may not be related to a low probability, per se, but to the level of contrast with the more likely or ‘normal’ not confirmed expectation. This means that, not all low-probability outcomes are necessarily surprising, even if surprises are generally created by low-probability events. However, as reported by the authors, even if their findings are more related to people’s cognitive representations of surprise than an emotional experience, it highlights the differences between surprise ratings and probability estimates. Thus, despite the fact that surprise is typically considered to be created by low-probability outcomes, this does not mean that all low-probability outcomes are necessarily surprising.

Novelty and surprise play a relevant role in human behavior and have been studied either through mechanistic models of neural processing or by psychological constructs. Most often, if not always, surprise accompanies novelty and has often been defined as a reaction to novelty (e.g. Berlyne, 1960). Yet, the opposite is not always true. A common example of this is the car door that is found unlocked. This familiar event is not novel, but we are surprised when we find our car unlocked.

In neural processing, an observation is considered to be novel when its representation is not found, or is not similar to another one stored in memory (Barto et al., 2013). Here, *novelty perception* regards the hippocampal general role of detecting mismatches between expectation and experience (Ploghaus et al., 2000; Strange & Dolan, 2001), in which the hippocampal neuronal activity represents the expected information or novelty of an event before it occurs (Strange et al., 2005).

Surprise is not a consequence of the attributes of luck or a random result, but rather an affective reaction to unexpectedness that stimulates causal thinking (Stiensmeier et al., 1995). It is unique to a particular event and is a specific consequence of the appraisal of novelty (Finkenauer et al., 1998), i.e., it measures its improbability or novelty (Strange et al., 2005). The relationship between the expected and the obtained result is what more strongly determines the surprise, which combines a previous experience and knowledge with the unfamiliarity of the outcome (Teigen & Keren, 2003). The overlap between both surprise and novelty has been observed too. Its result is usually related to attention capturing and learning (Ranganath & Rainer, 2003).

Surprise is generally accepted as an emotion that arises from the mismatch between expectation and what is actually observed, i.e., between an input coming from the outside and the individual's own schema (Ekman & Davidson, 1994; Derbaix & Vanhamme, 2003; Casati & Pasquinelli, 2007). Hence, surprise is elicited when the prediction based on an expectation is violated or frustrated (Bruner, 1986; Davison, 2004; Barto et al., 2013).

Furthermore, considering the concept of changing prior beliefs into posterior beliefs (Itti & Baldi, 2005, 2006, 2009; Baldi & Itti, 2011), surprise can be measured based, firstly, on the definition introduced by Shannon (1948)[10], secondly, by the probabilistic interpretation of an event given by the Bayes theorem[11], and thirdly, attending to the perceived emotion by the receiver. The empirical work developed in this dissertation relies on the latter, as will be described in the following chapters.

[10] The amount of information contained in a piece of data "D" is given by the probabilistic result of log2 P(D) bits and so, related to its rarity and small probability.
[11] It quantifies the amount of information included in a piece of data "D".

Additionally, particular attention will be given to the fact that when there is a novelty perception the emotion of surprise will follow. This will be particularly relevant when the use of surprise as a proxy of novelty is discussed to study the assumptions of bridging factors (Granovetter, 1973; Burt, 1992). The process of detecting novelty is not an easy task in a survey environment and surprise should be used instead. Furthermore, it is important to highlight the association between the cognitive process of appraisal (of novelty) and access to contents (i.e., interpreting). Interpreting can be seen as an output of meaning construction, enriched by the use of both conscious and unconscious processes of thinking. This is true particularly due to the unconscious processing given the associations between surprise, unconscious processing and construction of meaning. Thus, the use of surprise as proxy is also used to discuss the enrichment of construction meaning, as well as to understand the cognitive factors and constraints associated to information personalization on Web-based services. In this particular, the study by Flavián-Blanco et al. (2013) shows that online searching tasks have a positive impact on the positive emotions experienced after the search process. This is particularly related to feeling of "hope" that is usually satisfied when users find the information they have been looking for (or at least they perceive so). On the other hand, the lack of satisfactory results or process failure quickly originates negative emotions associated with the feelings of regret and frustration. Nonetheless, other cognitive factors are associated with this output, which we will be discussing in the two following chapters.

CHAPTER 3

WEB-BASED SERVICES AND COGNITIVE FACTORS: FROM A SEMANTIC WEB TO A SOCIALLY CENTERED WEB

3.1 Overview

The early stages of the World Wide Web's[12] evolution were characterized by institutions and companies offering information contents and services. However, with the shift towards a social-web (“Web 2.0”)[13], it started to be organized more around the users (Mislov et al., 2007) by means of social tagging (e.g., Kim et al., 2009; Pancke et al., 2009; Huang et al., 2009; Han et al., 2009), user-centric publishing and knowledge management platforms (e.g., Wikis and Blogs), as well as social resource sharing tools (e.g., Flickr), and social networks services (e.g., Facebook, LinkedIn, Twitter, Instagram)

In this second evolutionary stage, social networks have become an important source of data and information sharing among Web users. This shift as several implications, but the one that we want to emphasize is concerned with how a new way to access content was established through people’s social ties. Since then, the access is no longer through the common explicit links of the Web (Mislove et al., 2006), but via a social link that also carries implicit information that has become accessible and profitable for computation. Thus, when people are sharing contents on their social networks, they also are recommending information weighted by implicit information on the individuals interacting. Social networks are storing information after it is "filtered" by their members, and they serve also as a vehicle that connects thoughts. In this

---

[12] It is necessary to understand the differences between the Internet and the Web, terms that are often used interchangeably. The Internet is the physical layer or network made up of switches, routers, and other equipment. Its primary function is to transport information from one point to another quickly, reliably, and securely. The web, on the other hand, is an application layer that operates on top of the Internet. Its primary role is to provide an interface that makes the information flowing across the Internet usable. In this dissertation we use either one or other term, but always meaning the concept of Web. *In:* http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.

[13] Web 2.0 is the term used to define a computing paradigm that uses the Web as the application platform and facilitates collaboration and information sharing between users. See: O’Reilly, T.. What is Web 2.0: Design patterns and business models for the next generation of software. O’Reilly Media, Inc., (Sept. 30, 2005), *in*: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html.

context, we are witnessing the embodiment of thought through the content conveyed by the medium.

As mentioned by Press (1995), although Marshall McLuhan lived until 1980 and understood computers as a communication medium, he did not discuss it in his work Understanding Media, first published in 1964, or after its publication. However, when McLuhan (1964) claims that the use of new media was the prime cause of fundamental changes in society and the human psyche, it seems that he is projecting the same arguments as us, with regards to media effects and the intertwining cognitive factors and people' narratives.

Many saw Marshall McLuhan as a media thinker and visionary. He described at a very early stage of digital media, what life would be immersed in digital media narratives. As in my argument, the starting point for McLuhan is the individual, which entails a psychological dimension when media effects need to be thought about. As he defined it, media is as technological extension of the body[14].

We argue that the Web is emulating human narratives, which is reflected nowadays, for example, in the implicit information contained in the social links and in content interpreting which is shared.

McLuhan (1964) sustained that the electronic media are 'extensions of our nerves' (p. 152), and supplementary to this, we may add that in the scope of digital media, the Web is a representation of the extensions of meaning built by individuals, as a way in which current human thought is expressed and emotions are elicited according to content accessed and sender.

In fact, McLuhan's approach reflects a society where linear thinking (systematic thinking style) is extended to the electronic media. However, as mentioned by Press (1995, p. 16), "linear thinking may not be as important tomorrow as it was yesterday", given that tomorrow in 1995, is today. Hence, we conjecture that the McLuhan's 'extensions' could be seen today as the bridge between the collective consciousness (relative to explicit information) and the collective unconsciousness (relative to implicit information). With the concepts of 'collective conscious' and 'collective unconscious' being those proposed by Carl Jung (1959) and mentioned by Jones (2003).

[14] Biography: http://www.marshallmcluhan.com/biography/.

In this sense, I can reinforce my argument that we are witnessing the embodiment of thought through the medium. Through the medium we all are collectively sharing the broad amplitude of the construction of meaning, with its conscious and unconscious components. In fact, “What before was a mental process, a uniquely individual state, now became part of a public sphere. [...] Interactive computer media perfectly fits this trend to externalize and objectify the mind’s operations” (Manovich 2001, p. 60).

Thus, the new way to access content, carrying information contained in the social links (e.g., Mislove et al., 2006; Boyd, 2008), seems to recreate the relevance of a hyperlink structure as a path to content. It adds a new sense to that content and its path. Content is now associated to the individual who introduced it, as well as to individuals who will explicitly recommend the content. Thus, with the rise of publishing (by content creators that make information available to other users) and locating information (mechanism by which users find information relevant to them) through online social networks, search engines and recommender systems have found a new way to present customized results centered on the user. Furthermore, while, initially, the criteria of “individualized” and “interesting and useful” created a distinction between recommender system and search engines, nowadays, with the rise of personalization in information retrieval, this distinction is no longer visible (Burke, 2002; Adomavicius and Tuzhilin, 2011). The blurring of this distinction is even more present when the methods applied to personalization are based on “social” data gathered from Web 2.0 applications.

Accordingly, online social networks have not only become the epicenter of information sharing for many Web users, but also an important resource for the personalization and improving of several Web-based services. And similarly to what has been happening with the Web search, data from online social networks has also become tested on the improvement of recommendation.

Furthermore, although recommender systems have been comprehensively analyzed by scholars, the emergence of online social networks and the access to its data has sparked the rise of social-based recommender systems as a new and alternative method.

Through the evolution of those Web technologies, it was observed that some proposals, like Semantic Web and some social-web tools (e.g., tagging), have lost relevance or been abandoned by users. As an example, because tags are handled in a purely syntactical way, it means that the annotations provided by users create a very wide and noisy tag space that limits the effectiveness for complex tasks using this semantic approach (Lops et al., 2013).

Among other possible causes, the following sections present a review of such evolution and a trade-off between the evolution of some of these Web-based services and cognitive factors that may have justified its failure or abandonment. I focus on the cognitive factors to justify my argument.

In this vein, we question the current solutions of personalization based on social data. The use of these data are creating new technological challenges, whose opportunities also become threats due to the cognitive factors involved, more specifically, the lack of diversity in viewpoints and novelty. This is determined by the social organizing principles that justify the strength of the ties and similarities between friends. Further, this family of constraints arises from social data use – when based on strong ties and similarities among users – that motivates the research undertaken in this thesis, in particular on recommender systems.

Four main sections are presented in this chapter. First, is the *Semantic Web*, which reviews the challenge posed by a Semantic Web to inter-link the Web of contents automatically, and how differences in human construction of meaning end up being the main restriction to the success of a semantic-based tool. The second section, *Search engines and social network data*, reviews some solutions proposed by search engines and presents the boost of online social networks and how the information contained in the social links has recreated the sense of hyperlink. Cognitive factors identified by the term “Filter Bubble”, are presented here as severe constraints to innovation and search performance, notably from the users' viewpoint, given the current paradigm of social data use. The third section, named *Recommender systems and social network data*, presents the three main methods that have been widely debated and reviewed in the literature and in particular, the solutions that social data applies to improve recommendation. In fact, the latter use of information based on user behavior, similarities and social ties creates a new opportunity to present personalized recommendations and solve some persistent problems known in these systems, e.g., missing values of the user-item matrix. However, this approach, which improves recommendation, also presents some constraints. They are related to the familiarity of the recommendations provided to the target-user. In order to discuss this subject, this last section is organized into five topics: *Collaborative filtering methods; Content-based filtering methods; Hybrid systems; Social network-based recommender system; Novelty and diversity in recommendation;* and *Surprise in recommendation.* This last chapter ends by introducing the problem of Social Echo Chamber Effect, which we assume to be related to current use of social data, notably in social-based recommender systems.

3.2 Semantic Web

In 2001, referring to Berners-Lee's (1999) work, Ding (2001) mentioned that the World Wide Web was living a new technological shift with the Semantic Web. This new Web was providing additional automated services based on machine-processable semantics of data and heuristics using the metadata of the Semantic Web.

In 2010, the same author (Ding, 2010) contends that data can be represented with widely different syntaxes and semantics, which may make the task of integrating data very complex. This describes a reality that seems to be more complex than it was anticipated. However, the difficulties faced were not due to a lack of research. By analyzing the data from Scopus about the most cited articles on the topics of "semantics" and "ontologies" between 2005 and 2009, the author notes that the theme raised by the article of Berners-Lee (1999) still remains up-to-date ten years later. In fact, the number of citations found in the Scopus data base (Ding, 2010) illustrates the continuous investment made by the scientific community to achieve an automated Web, as was suggested in the initial vision of the founder of the WWW.

As is known today, in spite of the hard work of the scientific community, the problem has not been solved. This is possibly, because the solution is not rooted in technology, but rather in a more cognitive nature. This means that the difficulties of expansion and consolidation of an architecture based on the Semantic Web, machine-learning, or other automated services may lie not on the technological development but rather on the ability of the model to respond to cognitive factors. These cognitive factors seem to justify the unviability of the Semantic Web as a global technology, making the application of semantic solutions more restricted and applicable to contextually controlled environments, where the community is aware of the used vocabulary (Pollock, 2009) and of its boundaries of meanings.

In fact, it is commonly accepted that in terms of word reading activity, the meaning is not a property of words, because words have no meaning (Evans, 2006). The meaning of words is related to each individual and how he uses it in his own cultural context, knowledge and experience, which makes the meaning extracted from words and utterances flexible. Moreover, concepts related to cognitively irreducible key areas (linguistic) have a reduced interpretive ambiguity, but this is not the case of linguistic expressions, because they are not related to key areas (most of the vocabulary).

Additionally, given that the majority of the linguistic expressions do not refer to fundamental domains, but to higher levels of conceptual organization, each word can elicit an infinite number of cognitive domains, and so, be related to countless application contexts (Lévi, 1991). Thus, given the multiple meanings of words, when an individual is accessing his mental lexicon, he is accessing more meanings than others, depending on his/her cognitive priorities (Kecskes, 2006), which underlines the fact that the construction of meaning differs between individuals. That is why construction of meaning cannot be based only on the explicit information conveyed by language, since language doesn't provide a unique meaning or a true meaning (Freeman, 2003). It varies from person to person.

On the other hand, the need for compromise to create a unique ontology of meanings became one of the main bottlenecks of this technology and a difficult constraint that the ontologies have tackled. This was a compromise that needed to be achieved by all the interested parties. Otherwise, it would be an imposition of a vision (or several), but not representative of all visions. This vision means the understanding of the person (or persons) that design the ontologies. However, if it was imposed a solution not agreed by all interested parties, I could argue that the «Dewey's error»[15], by analogy, would be recreated. "Dewey's error" is based on the organization principles proposed, which are oriented through one single perspective and from a 'physical' way,

[15] In 1876, John Dewey proposed the Decimal Classification System basing it on his understanding of what knowledge was. His proposal was based on a representation of his "physical" world (Weinberger, 2007). The proposed system can be regarded as a second-order way of organizing the information, which is constrained by the physical reality of paper and the need to give each book a single spot on a shelf (Weinberger, 2007).

instead of being based on the 'meaning' that combines different perspectives from different people. It is a systems based on the proposal from one to all.

In short, in order to unify the proposed technology, the implementation of standards namely at language level is difficult and highly controversial. Thus, the inevitable lack of convergence on meaning construction given the richness of language and cognitive differences between people, notably at their unconscious level, characterizes some of the constraints related to semantic-based solutions. Hence, the afore-mentioned cognitive factors should be highlighted as the main drawbacks for the enhancement of this technology. Therefore, the cognitive factors underlying the interpretation of semiotic signs significantly increase the ambiguity of this technological proposal.

Furthermore, they draw our attention to the challenges that semantic tools may present when considering an online large-scale use, i.e., beyond restricted environments at vocabulary level and contextually controlled.

3.3 Search engines and social network data

Web search engines have transformed the way people find, share and perceive information. Recent studies have shown that searching for information, together with email services, is the most frequently performed activity by users (AECE, 2009)[16]. In fact, one of the main informational retrieval tools that users have at their disposal is the online search engine (Rangaswamy et al., 2009), of which Google is the most visited website in the world (Alexa, 2014)[17].

It can be said that the semantics of a search engine are "matching", since it is supposed that the system returns the items that match the query ranked by degree of match. Two main developments can be highlighted for information retrieval systems or search engines: a) the "authoritativeness" criteria incorporated by Google (Burke, 2002); b) Personalized social search.

a) Online search evolved dramatically when Google incorporated the "authoritativeness" criteria (Burke, 2002) into its ranking (defined recursively as the

[16] Asociación Española de Comercio Electrónico y Marketing Relational – AECE, ''Estudio sobre comercio electrónico b2c 2009". (2009). http://www.red.es/media/registrados/2009-10/1256816746333.pdf?aceptacion=8686d2aacf93732ad9c39ce7ba5f0018/..

[17] Alexa, "Top 500 global sites". (2014). http://www.alexa.com/topsites/ Retrieved July 2014.

sum of the authoritativeness of pages linking to a given page) aiming to return more useful results (Brin & Page, 1998).

The hyperlinks between content (typically pages) form a hyperlink structure based on the incidence of links to Web pages. This is the primary tool for structuring information. It is this structure of links that informs search engines of the corpus of information to be indexed, i.e., to crawl the Web to index content. Hyperlinks also inform the search engines about the relevance of a certain Web page relative to a given query. It allows estimating and ranking the relevance of the content (Page et al., 1998; Mislove et al., 2006).

On the other hand, hyperlinks identified as explicit links are also used by people as an indicator of relevance of the browsed content, as well as to embed a Web page in the context of related information (Mislove et al., 2006).

b) Despite continuing improvements in this hyperlink-based search paradigm, some limitations were reported in literature (e.g., Mislove, 2006), highlighting two main concerns: i. how to make a new Web page or content for the search engine visible; and ii. how to avoid the biases introduced by the incident link solution to rank the importance of a certain Web page or content.

At the center of these concerns is the meaning of "user" and how to integrate the "user" into studies of information retrieval (e.g., Jones, 1988). Rather than continuing research almost exclusively on document representation (Belkin, 2008), it was found that we needed to approach the meaning of user in a different way. Nonetheless, it was necessary to form a more consensual understanding of this need to react positively to the questions raised by Sparck Jones in 2008. After that and benefiting from the amount of social data available, scholars started to develop a more personalized social search. With that, the information exchanged in online social networks started to be examined as a source of naturalistic behavioral data.

This type of search requires the ability to model the users' preferences and interests – done through the tracking and aggregation of users' interaction with the system (Carmel et al., 2009). Some examples of user aggregation are represented by tracking information on users' previous queries (Tan et al., 2006), or click-through analysis (Dou et al., 2007). The interaction of the users with the system can be represented by users' profiles that are applied in the search (Agichtein et al., 2006). This can be employed by incorporating users' interests with the processes of re-ranking and filtering of the search results (Shen et al., 2005). However, this approach may raise issues related to privacy, because user profiling may be understood as a violation of user privacy (Carmel et al., 2009).

The concept of social search has several alternative definitions. We adopt the one used by Carmel et al. (2009), which states that social searches "describe the search process over "social" data gathered from Web 2.0 applications, such as social bookmarking systems, wikis, blogs, forums, social network sites (SNSs), and many others" (p. 1228). Thus, the explicit user interactions provide an ideal framework for personalization. The assumption behind the use of data based on a user's social network to obtain user preferences from related people is that closely related people have similar interests[18].

Facebook presents a relevant example of the trade-off between service provided and information gathered from users, in order to apply it to other Web-services (e.g., on Bing[19]). While Facebook gives to its users the opportunity to make their relationships explicit (among friends and acquaintances) and share all kinds of information, the algorithm infers intimate details about users' preferences. Applying these data (carefully, given privacy concerns) can improve greatly other Web-based services, like searching.

In this regard, Piscitelli et al. (2010) states that filtered content based on our social network is likely to provide information that is equally or even more relevant than content obtained through standard Web searches.

Several approaches for directly or indirectly employing users' social relations to improve personalization have been proposed by scholars.

Mislove et al. (2006) tested the use of social network information to inform and bias the ranking algorithm of a search engine and found an improvement of 9% in search result clicks over Google alone. This integration has the potential to improve the

---

[18] Similarly, this is one of the main assumptions behind collaborative filtering methods in recommender systems.

[19] Greene, J., 2012. Bing deepens Facebook integration, connecting searchers with friends. From http://www.cnet.com/news/bing-deepens-facebook-integration-connecting-searchers-with-friends/

quality of Web search experience, because nearby users in the network often find relevance on similar sets of pages.

Golbeck et al. (2007) proposed the integration of social network information into a user's browsing experience using a Firefox extension. The goal was to create additional contextual information loaded from other sites on the subject browsed by the user. This means completing contextual information with data on what others are saying about the subject browsed.

Bender at al. (2008) developed a framework representing a social community by means of a network graph model of users, documents, and user-generated annotation (*tags*) that gives information about users' interests and users themselves. They found that social expansion based on the Friendship graphs improves the precision of the retrieval effectiveness remarkably. These results contribute positively to social search strategies.

Carmel et al. (2009) analyzed the value of personalization according to different relationship types, in particular familiarity and similarity. The results show that social network based personalization significantly outperforms non-personalized social search.

Cai et al. (2014) raise the question of search engines' lack of ability to be aware of users' interests or how to efficiently find the information that users need. So, the authors propose to store users' search history in the user profile and relocate the results of search history by the particular subject. The proposed method provides a personalized search service that gives priority to the documentation already seen by the user to position it at the top of the search results.

Several other contributions regarding personalized search have been presented in literature, such as Song et al. (2014), who adapted the well-known ranking model of RankNet to personalized search, or Gasparetti et al. (2014), who selectively collect text information based on implicit signals captured through web browsing interactions of the user.

However, privacy concerns a side (Carmel et al., 2009; Younus et al., 2014) that is not free of cognitive constraints. Thus, here too, there is setbacks on innovation and performance of information technology, which can be pointed out to cognitive factors. This constraint is described in literature as Filter Bubble.

Filter Bubble

The term *Filter Bubble* was introduced by the internet activist Eli Pariser[20] to describe how algorithms are tailoring information to people, creating a personal ecosystem of information based on user information (such as location, past click behavior and search history). As a result, users are separated from diversity and receive what 'can be' expected. In this sense, personalization confirms what people already know and avoids offering information that disagrees with the user's viewpoints. In short, the term *Filter Bubble* describes the potential for online personalization to effectively isolate people from a diversity of viewpoints or content (Nguyen et al., 2014).

This effect has been most noticed in Google results since December 2009, when it started to customize its search results. Several bloggers have already stated their concerns. An example of that is Cyrus Shepard[21], who contends that:

> "[personalisation] creates a real risk of limiting our worldview. Every new search result starts to look like the search before. Our ideas become isolated and homogenized, like exclusively watching only Fox News or MSNBC, while refusing to consider CNN. There are times when personalization and localization work well, such as when I'm looking for a pizza restaurant in Seattle. The maddening part is, what if I want to turn it off? There are times when I want unbiased results not based on my past search history, my location, or what my social circle has shared."

Other bloggers[22] used the term "Echo Chamber Effect" to refer to the problem described above as the "Filter Bubble". The constraints underlined are the same and centered on the problem of personalized search. The difference is that this is an individual option, referring to the fact that people freely decide their "political corner", for example, as expressed by Jamieson & Cappella (2008) to explain the homogeneity among people that share similar political views. This kind of personalization / forced homogenization is determined by a 'blind' algorithm, which is socially ignorant and,

[20] http://www.amazon.com/The-Filter-Bubble-Personalized-Changing/dp/0143121235.
[21] Cyrus Shepard: http://moz.com/blog/google-personalized-search.
[22] Grant Jacobs: http://sciblogs.co.nz/code-for-life/2011/07/30/google-and-the-echo-chamber-effect/.

probably, biased by business interests (e.g., some sort of tailored advertising). This problem is not only affecting online searching though, but also other media, such as Facebook (e.g., through the *news feeds*), which is tailoring the results based on personalization.

Nowadays, we are still witnessing the continuous growth in access to social data, which has increased the availability of more and more explicit and implicit information. This fact is yielding even more insight to develop new digital media solutions, notably in the improvement of recommendation through recommender systems. However, problems related to cognitive factors persist.

3.4 Recommender systems

Web entrepreneurs at the forefront of the information revolution were the first to notice the opportunity that recommendation could leverage. This is one of the reasons why the study of recommender systems is at the intersection of science with business, i.e., they are an integral part of some e-commerce sites (Schafer et al., 1999), calling for contributions from diverse knowledge fields, such as computer scientists, mathematicians, physicists, and even psychologists and sociologists (Lü et al., 2012).

A recommender system suggests items of interest to users based on their explicit and implicit preferences, given the preferences expressed by other users, and attributes of the users and items. A recommender system is expected to predict users' possible future likes and interests based on data from the users and their preferences. The main basis is the act of suggesting items based on a representation of what a user likes and dislikes, with the aim to personalize, as much as possible, the delivery of the right content to the right person.

The recommendation activity is particularly relevant for sales based on the so-called long-tail (Anderson, 2006). The long-tail refers to goods that are rarely purchased, but given their multivariate they represent, in total, a great quantity, and so they can yield considerable profits for the businesses able to explore this model (Leskovec & Adamic, 2007). This is the case of *Amazon.com*, where 20 to 40 percent of its sales are based on products that are above the line of the 100 000 most sold products (Brynjolfsson, 2003). Another typical activity of recommendation is the sale of goods, like DVDs rented by Netflix. Here, the purchases based on personalized recommendation achieved 60 percent in 2009 (Lü et al., 2012).

Three main methods have been widely debated and reviewed in literature (e.g. Bilgic, 2004; Adomavicius & Tuzhilin, 2005; Lü et al., 2012; Park et al., 2012), in which authors present extensive surveys with the pros and cons of each system and suggestions for new solutions. Next, we present an overview of these methods.

3.4.1 Collaborative filtering methods

Collaborative filtering (CF) (Adomavicius & Tuzhilin, 2005; Shi et al., 2014) and content based filtering (Chen & Sycara, 1998) are the most common types of recommender systems, and hybrid systems combine the strengths of both types of systems (Burke, 2002). CF is the most successful method and can be found in several online applications and fields of knowledge, such as health education (Luque et al., 2009), consumer reviews (e.g., Epinions.com) (Massa & Avesani, 2007), and sentiment prediction in twitter conversation threads (Kim et al., 2013). New approaches have also been developed, like the one introduced by Cai et al. (2014), who propose a user-based recommendation that makes a representation of the user through a vector that can indicate the user's preferences on each kind of item. It finds a user's neighbors based on their typicality degrees in all user groups. This is different than rely on users' ratings on items as happens in other methods.

Apart from some recent proposals, like Cai et al. (2014), two common types of CF have been discussed in literature: a) user-based recommendation, and b) item-to-item recommendation (the Amazon model).

a) In the user-based recommendation, the system finds similar users (collaborative) and makes a prediction based on those similar user preferences (filtering) (e.g., Ali & van Stam, 2004, Arora et al., 2014). The principle is to pick people who share similar tastes with someone else, and make an automatic prediction about the taste of someone based on the collected information from many others. It can be summarized by the idea that "People like you bought, liked or shared Y". In this regard, Adomavicius & Tuzhilin (2005) present the well-known user similarity method, based on the taste overlap between users. This technique recommends items frequently collected by a given user's "taste mates".

b) In the item-to-item recommendation, the items are compared first, but incorporating user preferences. It takes the preferences of users who liked (or bought) one item to suggest an item those users liked just as much (this system was made popular by Amazon) (Schafer et al., 2001; Linden et al., 2003; Koenigstein & Koren, 2013). The idea that summarizes it is: “People who bought, liked or shared X also bought, liked or shared Y”.

CF recommendation presumes that people who have similar tastes will rate items similarly. This method bases its recommendations on community preferences (e.g., user ratings and purchase histories), ignoring user and item attributes (e.g., demographics and product descriptions) (Schein et al., 2002). Thus, to predict a recommendation about a consumer item (e.g. a book, a film) the item needs to have a reasonable number of ratings. However, this is not always the case, particularly for new entries (new goods). Similar constraints occur when the target user (recommendee) has unusual tastes compared to the rest of the population that has evaluated the items, which makes it even more difficult to find a similar profile. Thus, all of these complications lead to poor recommendations. These constraints are known as rating sparsity.

The problem of sparsity data occurs when there are several items to be recommended and the user/ratings matrix is sparse, independently of the number of users, which makes it difficult to find users that have rated the same items (e.g., when someone bought only one book on Amazon it is hard to accurately determine similar preferences, given the lack of information on the user and few overlapping items).

To overcome the problem of ratings sparsity, some scholars have been exploring solutions based on demographic information, known as "demographic filtering" (Pazzani, 1999), transitive trust graphs, as a way to increase the number of comparable users (Massa & Avesani, 2007), social information (Kaya & Alpaslan, 2010), or even by the selection of optimal personal propensity variables (Jeong et al., 2013), just to mention a few of them.

In the case of Massa & Avesani (2007), the goal was to search for trustable users in a social trust network, instead of searching for similar users in a social network (e.g., friends of friends). The social trust network is based on user feedback about which recommendations they trust most. This feedback is used to rank people in the trust network.

Other examples incorporating trust network into CF are presented in Yan et al. (2013) and Gou et al. (2014). Yan et al. aims to resolve the neighbor selection problem, while the Gou et al. proposal distinguishes between implicit and explicit trust. Here, it is argued that the inference of trust based only on user ratings is not sufficient to capture the dynamics and context-dependency. The authors suggest the incorporation of contextual information when the ratings are given, as well as the users' interactions pertaining to the items.

Some other challenges in CF are the cold-start problem (e.g., when new users have zero ratings/purchases, or items that no one – in the data set – has yet rated) and the popularity bias (e.g., everyone reads "Harry Potter", or someone with unique tastes).

Finally, given that CF bases its recommendation on overlap, i.e., similarities, rather than differences, this narrows the access to novelty and to different viewpoints, by exposing them, mostly, to a narrowing band of popular objects. As a result of this, a niche of items that might be very relevant will be overlooked (Zhou et al., 2010), which will emphasize the problem of the Filter Bubble, as studied in Nguyen et al. (2014). The authors conclude that there are two forms that represent the narrowing of influence of an online recommender system on its users: a) through the items recommended by the system, and b) the items rated by users. Furthermore, the risk of a filter bubble increases when users follow recommendations that appear in their top-N recommendation lists. As a matter of fact, as Ziegler et al. (2005) show, user satisfaction can be improved with diversification.

3.4.2 Content-based filtering methods

Content-based filtering methods rely on comparing content of items rather than on other users' opinions. It uses an algorithm to induce a profile of the user's preferences from previously rated items, matching query words or other user data with item attribute information (Mooney & Roy, 2000). The goal is to recommend items that fit this preference profile based on similar content. In this sense, some authors defined the design of similarities from an inter-concept similarity based on the distance of the concepts to their least common subsumer in an ontology (Fernández et al., 2006). This solution presents at least two variations. One applies taxonomies as a basis for calculating similarity, and the other uses only annotated corpus data. Here it is the frequency with which concepts are used that defines similarity (e.g., Lin, 1998).

Nonetheless, these solutions found similar constraints to those faced by the ontologies proposed in the Semantic Web. This limitation resides in the fact that the filter does not distinguish between word senses (Tintarev & Masthoff, 2006).

Some of the known constraints related to that approach are the lack of scalability and the fact that it is not social. The advantage is that there is no need for data on other users, which avoids the cold-start and sparsity problems. Also it is able to make recommendations to users with unique tastes, as well as recommending new and unpopular items, which avoids the so-called first-rater problem. Moreover, it can provide explanations of recommended items by listing content-features that caused an item to be recommended.

An example of a socio-economic application of this method is Pandora.com, which is a free and personalized radio that plays music online. Pandora coded the so-called "genome"[23] of each song to generate personalized recommendations based on “genes” from songs that users liked. The Music Genome Project (Liu et al., 2009) attempts to analyze the content attributes of each song. Based on the name of songs or artists typed by users, the system finds requests which are similar to make recommendations. Another example of a content-based approach, this time applied to the culinary domain in recipe recommendation, is proposed by Lin et al. (2014).

However, there are also challenges with this type of recommender system which need to be overcome. First, it requires content that can be encoded as meaningful features. Second, users’ tastes must be represented as a learnable function of these content features. Finally, it is unable to exploit quality judgments of other users, unless these are somehow included in the content features recommended.

3.4.3 Hybrid systems

Hybrid systems combine collaborative methods with content-based methods or with different variations of other collaborative methods. This method is helpful to address the diverse needs of heterogeneous users (Burke, 2002), or to join the best of different methods, as in Lops et al. (2013), who proposes a tag recommender system implementing both a collaborative and a content-based recommendation technique. For example, CF is useless in solving the problem in a cold-start setting, but content information can help to bridge the gap from existing items to new items by inferring

[23] In Wikipedia: “The Music Genome Project is an effort to "capture the essence of music at the most fundamental level" using almost 400 attributes to describe songs and a complex mathematical algorithm to organize them. The Music Genome Project is currently made up of 5 sub-genomes: Pop/Rock, Hip-Hop/Electronica, Jazz, World Music, and Classical. Under the direction of Nolan Gasser and a team of musicological experts, the initial attributes were later refined and extended.”

similarities among them. Hybrid systems that join CF and content-based methods enable us to solve this problem (Schein et al., 2002; Tewari et al., 2014), which is one of the most important applications of this type of recommendation system.

Some proposals for hybrid systems aim to recommend long-tail items, to which users have had little access – as seen above (Zhou et al., 2010). This method provides novel insights about users by combining collaborative filtering with graph spreading techniques.

3.5 Recommender systems and social network data

3.5.1 Social network-based recommender system

After being strongly focused on how algorithms could better predict unrated items, scholars started to look more carefully at solutions based on social elements of decision making and advice seeking (Bonhard, 2004), by merging recommendation systems and social networks (for a review see Bobadilla et al., 2013; Tavakolifard & Almeroth, 2012).

In this vein, scholars found that people seem to appreciate more a recommendation coming from friends than one from a recommender system (Sinha and Swearingen, 2001), which underlines the relevance of a recommendation based on social-influence in relation to similarities of past activities. As mentioned by Lü et al. (2012), scholars have understood the value of social influences for a long time, yet, it was with the emergence of Internet and particularly with the rise of social networks that it has become possible to understand social influences quantitatively.

The effects of social influences can be divided into two classes: a) users' prior expectations, which lead to the increase of sales, and b) users' posterior evaluations, connected to the improvement of user loyalty.

a) Leskovec & Adamic (2007) studied the effects of social influences on purchase preference in an e-commerce system. The authors tested the reaction of the target users to recommendations from friends through e-mails after purchases. The results reveal that individuals are often impervious to the recommendations of their friends, particularly when the recommendations received arrive at saturation level (about 10 recommendations for DVD products). At the same time, with book sales, they reported that the purchase probability had little effect or was even negatively affected after the recommendations. In short, the authors concluded that there are limits to how influential high degree nodes are in the recommendation network. There is a limited reach of influence that individuals have over friends – they just reach a few of them, and, furthermore, they do not reach everybody they know.

b) In order to solve traditional challenges related to CF, such as data sparsity and cold-start, and harnessing the emergence of online social networks, some authors merged both areas to enhance a social network-based recommender system. Social recommendation introduces transparency to the activity of recommendation and a higher level of trust in the system itself (Groh & Ehmig, 2007).

Lü et al. (2012) analyze how to utilize social network information in social recommender systems. The authors present a framework incorporating social context information and show that this improves the accuracy of review quality prediction, in particular when the data is sparse.

He et al. (2010) crawled the dataset of the online social network Yelp.com to analyze whether or not friends tend to select the same item, and whether or not friends tend to give similar ratings. The results reveal that friends have a tendency to review the same restaurants and give similar ratings. Also, on the sparsity test and cold-start test, the proposed system performs better than on CF.

Ma et al. (2011) provide a general method for improving recommender systems by incorporating social network information, in particular, to solve the persistent problem related to the missing values in the user-item matrix. The results are indicative of improvements in recommendation, but by using all the social connections of each user, the recommendation performance decreases. The authors do not explain the reasons why this happens, but propose for future work the use of an algorithm that identifies the most suitable group of friends for different recommendation tasks.

There is one thing the approaches mentioned have in common, which is that they all lead to some kind of challenge that researchers in the field of recommender systems have to face. Some of the major challenges have been identified as sparsity, scalability, cold start, diversity vs. accuracy, vulnerability to attacks, etc. (for a complete review see: Lü et al., 2012), while other issues have been pointed out by scholars, particularly the danger of an excess of recommendations based on popularity (Zhou et al., 2010), low novelty, or lack of diversity (Vargas & Castells, 2011). In fact, just as happens in

Web searches, diversification of results is a critical factor in significantly influencing user satisfaction[24].

In this regard, Zhou et al. (2010) estimated the capacity of a recommender system to generate novel and unexpected results by measuring the unexpectedness of an object, i.e., the average self-information or “surprisal” (term coined by Tribus, 1961) of recommended items, which amounts to the average log inverse ratio of users who like the item (also known as “inverse user frequency”).

Another issue on recommendation is the absence of control of number of times that the system recommends the same items to users over and over again, or whether a novel content is delivered in recommendations, or to the appropriate user.

In this scope, Lathia et al. (2010) studied the novelty that a system delivers with respect to recommendations that it produced in the past. They observed that CF algorithms often repeatedly recommend the same (top-N) items to users. To invert this, the authors suggest switching the CF algorithm over time, in order to re-rank the results of frequent visitors to the system, making that a temporally evolving system that could give diversified recommendations in time. In turn, Vargas and Castells (2011) proposed an evaluation of the novelty and diversity of the recommendations attempting to formally unify them in a single evaluation framework.

Therefore, the familiarity of online recommendations characterized by the Portfolio Effect concept (Burke, 2002) is then a problem addressed by scholars, although, not from the viewpoint of a cognitive constraint increased by the current use of social data. In order to study complementary solutions and still benefiting from the richness of social data, it is important to discuss the role of the emotion of surprise, given its relationship with novelty.

3.5.2 Surprise on recommendation

Surprise in recommender systems can be observed in several ways. I present three approaches: serendipity, past surprise, and network approach.

[24] Diversity has been addressed by scholars in the scope of personalization of Web searching and with promising results (Vallet & Castells, 2012).

The relevance of surprise from the viewpoint of users' emotional reaction is related to the process of information sharing/ content selection debated in this dissertation.

The serendipity and past surprise approaches, because they have often been used by scholars and entrepreneurs in order to trigger the emotion of surprise on the online users when they access the contents.

In this dissertation it is explored the network approach. This is a not yet explored approach that uses surprise as a proxy of novelty – surprise response is elicited in the context of a network bridge (Granovetter, 1973; Burt, 1992). This subject is detailed in the next chapters.

To explain serendipity, some literature mentions the process of incidental information acquisition (IIA) as an occurrence in which a person acquires information (useful or interesting), while not consciously looking for it (Williamson, 1998; Heinström, 2006). This acquisition is due to the individual's psychological receptivity that makes people more or less attentive to the received message (Heinström, 2006). Thus, personal traits and emotional states may determine the attention to the message.

Some studies (Swearingen and Sinha 2001) recognize that "surprise" caused by serendipity in recommendation expands their horizons, while others (Groh & Ehmig, 2007) say that serendipity can convey novel predictions in recommendation, which are brought from cliques (clusters or groups of people that share similar tastes). This same approach is argued in Zhang et al. (2012), whom propose a framework of novel recommendations based on serendipitous recommendations.

The past surprise approach, on the other hand, is explored by Horvitz (2007). The author explains the "mixed-initiative interaction", which allows the collaboration between computers and humans in which human skills will attempt to expand the ability of the computational systems. The author[25] believes that through the technique of 'surprise modelling', it would be possible and beneficial to consider the kinds of things that have surprised people in the past to model the kinds of things that may surprise them in the future.

The third approach results from the information flow in social structures situating surprise as an emotion which arises when the input coming from the

[25] http://www.technologyreview.com/read_article.aspx?ch=specialsections&sc=emerging08&id=20243

surroundings doesn’t match the individual’s own schema (Derbaix, 2003). Surprise can be either positive or negative and can be related to different types of communication processes, e.g., action tendency of "interrupting" (Frijda, 2003), cognitive dissonance (Festinger, 1957).

Considering the situation of communication in which the sender and the receiver of information are connected by a network bridge, the surprise might not only be caused by the information received, but also by the perspective delivered within the information. This is to say that the cause of surprise might not only be the information itself, but also the sender of information associated with structural and relational properties and individual attributes – characterizing psychological characteristics. In the following chapters it will be presented greater detail on this.

Given this, in the context of this dissertation, surprise must not be elicited in a serendipitous way, because a system that relies on this kind of design cannot compute this emotional state from data based on naturalistic behavior from social networks. On the other hand, through solutions such as the "mixed-initiative interaction," surprise relies on a probabilistic approach that may not fit the purpose of my approach.

There are many highly diversified contributions involving recommendation and social network data, numerous approaches to enhance novelty and diversity, as well as several methodologies to assess and measure how well this is achieved. Though, despite the common understanding that social influence and data on naturalistic behavior from social networks are very relevant to improve accuracy on recommender systems, it seems that there is an important concern in this regard that should be considered. This is related to network mechanisms supported by homophily (McPherson et al., 2001), like triadic closure, which contributes to closure and so to reducing novelty access.

On the other hand, if the priority or opportunity is on the use of social network data, then novelty can be approached from a social and psychological perspective too. Mastering both factors in the same framework, this might bring new insights about Web-based systems and particularly to recommender systems. This dissertation dedicates particular attention to this opportunity.

In this context, the following chapter introduces the concept of Social Echo Chamber Effect and the problem of trapping users inside their own social bubble – echo chamber – when the recommendation's target user receive a recommendation based on data from similar users or close friends (strong ties).

CHAPTER 4

SOCIAL ECHO CHAMBER AND SOCIAL STRUCTURE

4.1 Overview

This chapter introduces the concept of Social Echo Chamber Effect applied to Web interactions. The point of departure is the assumption that similarities in personal attributes (e.g., attitude, age, race, ethnicity, education, religion, socio-economic status, physical, etc.), between individuals who are socially connected, notably by strong ties, are associated to a low level of novelty exchanged and lack of diversity of viewpoints among them. I advocate that the use of data based on these kinds of attributes to improve the personalization of Web-based applications, rather than boosting innovation and opportunity, may generate the "Social Echo Chamber" effect. Although the problem of "Filter Bubble" or "Echo Chamber Effect" has already been identified in literature in the context of online personalization (e.g., Nguyen et al., 2014), in this work I study this problematic from the perspective of social data use. I emphasize that the use of this kind of data may help in personalization algorithms, but it also has some drawbacks, because it may generate dissatisfaction in users. This highlights that the relationship between cognitive factors and social interaction may distort the meaning of Web personalization. Thus, it is important to fully understand this phenomenon in order to find solutions.

This chapter is organized in three sections. The first section, named *Echo Chamber Effect*, introduces the concept of the "Echo Chamber Effect" and reviews the literature. The second section, called *The Social Effect on Echo Chamber,* shows how social dynamics based on homophily and strong ties contribute to the reduction of diversity of viewpoints among groups and thus provoke the effect of Echo Chamber. Four main topics are developed: *Novelty and diversity*, *Homophily on Echo Chamber*, *Strong ties on Echo Chamber* and *Network mechanisms on tie formation that feeds the echo chamber*. Finally, the third section called *Social Echo Chamber Effect on Web Personalization*, presents the effects of echo chamber on personalized recommendations explaining its relationship with the current use of social data.

4.2 Echo Chamber Effect

The term Echo Chamber Effect was coined to emphasize the human behavior that is typically observed in political or cultural communities (Jamieson & Cappella, 2008), whose individuals seek information or join groups similar to their prior beliefs and biases (Sunstein, 2001). This behavior leads to the argument that the look alike effect plays an important role on self-affirmation “birds of a feather flock together” (McPherson et al., 2001). A person typically enjoys receiving confirmation of every aspect of his or her ideas and attitudes. It is argued by some scholars that this kind of social interaction results from homophily, given individuals' preferences to interact with others that have similar background or opinions (McPherson et al., 2001). This can be explained either by the structural constraints of society, which limits people’s social worlds (Blau, 1980), i.e., their ability to interact with others from different backgrounds or with different opinions, or individual choices within social structures (McPherson et al., 2001), or even as a result of social influence through interactions (Ma et al., 2009).

Consequently, people in these situations lack exposure to diverse viewpoints. Social structures grounded in homophilous relationships and in strong ties set by friendship, have a higher likelihood of increased information access from other strong ties that keep spinning the same personal perspectives. Strong ties are then characterized by their homophily (McPherson et al., 2001), which strengthens the bond between people contributing for them to have the same close friends either online or offline (Lewis et al., 2008; Hampton et al., 2010).

The massive amount of media currently on offer would seem to ensure exposure to a broad spectrum of views and the diversity of information for a healthy democracy. However, that is not necessarily the case. First, as already reported by Centola (2007), people joining a new online group seek similar people, which maintain the same offline cultural affinities – i.e., the homophily effect is also visible in options made by people who select their communities online. Second, it has been reported that similar people present similar behaviors of information access. People look for content that, in fact, feeds their prior views, i.e., “people avoid the news and opinions that they don't want to hear”[26]. Consequently, this circular option locks the individual in an experience of echo chamber. On top of this, the current fragmentation of the communication market and the

[26] http://press.princeton.edu/titles/8468.html.

concentration of ownership seem to emphasize this problem, as sustained by Cass R. Sunstein in the book “Republic.com 2.0” (Sunstein, 2009). In this sense, people are "closed" inside their own cluster. A cluster that is naturally created and tuned by endogenous properties, i.e., homophily, as basic social organizing principle, but which is treated as a selected audience by the media. As a result, this audience becomes externally regulated by a pre-constructed and imposed view by the media. Furthermore, current social networking systems (SNSs) seems to emphasize narrowing contexts of information as detailed by Boyd (2002) through the concept of *collapsing of context*. It has been stated in this scope that the digital world alters people's notions of context and identity. Gilbert (2012, p. 2) posits in the scope of this concept that, “in social streams, people from every part of life collapse into one channel, in temporal order, with nothing distinguishing one from any other.” This author uses the following analogy to explain: “imagine living your whole life at your own wedding. Everyone you know from various parts of your life is there: grandmothers, in-laws, coworkers, cousins, childhood friends, etc. Writing a status update on a social media site is like forgetting you left the microphone on. Everyone hears everything. Consuming content (e.g., reading Twitter or the Newsfeed) is very much like standing in the receiving line. Everyone you know passes by in random order.” (p.2)

4.3 The Social Effect on Echo Chamber

Having established the conditions that are facilitating the emergence of the echo chamber factor, we will now discuss its effect when Web applications use data from supposedly like-minded people to personalize the information services required. In this context, it is important to understand what network dimensions are behind the social data in the Web personalization and why that is reducing diversity and novelty potential.

Novelty and diversity

As established earlier, the information based on interactions from strong ties has a very low rate of novelty (Granovetter, 1973). On the other hand, the concepts of novelty and diversity are intertwined, which means the argument that the low level of novelty in the information accessed leads to a deficiency of variety in the information shared, which may reduce the quality of the contribution of social data on recommender systems. This idea is underlined by Vargas & Castells (2011), which states that:

> "Novelty and diversity are different though related notions. The novelty of a piece of information generally refers to how different it is with respect to "what has been previously seen", by a specific user, or by a community as a whole. Diversity generally applies to a set of items, and is related to how different the items are with respect to each other. This is related to novelty in that when a set is diverse, each item is "novel" with respect to the rest of the set. Moreover, a system that promotes novel results tends to generate global diversity over time in the user experience". (p.2)

To clarify how network interactions may participate in keeping the echo and then in reducing the access to novelty, we need to develop further some properties related to homophily, tie strength, and network mechanisms.

Homophily on Echo Chamber

There are plenty of published social network studies on bridging factors (Granovetter, 1973; Burt, 1992) or centrality (Freeman, 1973) to explain outputs related to the information flow (e.g., McEvily et al., 1999; Hansen, 1999; Holme & Ghoshal, 2008, Kratzer & Lettl, 2008; Shi et al., 2013). In this regard, some authors argue that weak ties are more likely to be dissimilar than strong ties relatively to the ego ('owner' of a social network), and that this dissimilarity is advantageous to expose the ego to a dissimilar knowledge and new perspectives (Zhou et al., 2009) and so, to influence the information flow.

On the other hand, several socio-psychological studies indicate that homophily is a noticeable characteristic of social interactions despite often being diminished and attributed to peer-to-peer influence[27] (Aral et al., 2009), given the preference to interact with people with similar attitude, background or opinions (McPherson et al., 2001).

---

[27] Homophily has been seen to be more important than peer-to-peer influence and more relevant that sometimes is assumed by scholars, notably, in the sense that it can account for a great deal of what appears at first to be a contagious process (Aral et al., 2009). However, the distinction between

The term homophily coined by Lazarsfeld & Merton (1954) suggests that individuals tend to associate with others who share similar backgrounds or opinions and is often referred to as “similarity breeds connection” (McPherson et al., 2001). It can be said that homophily is "the conscious or unconscious tendency to associate with people who resemble us" (Christakis & Fowler, 2009, p.17).

The homophily between two individuals – as a tendency to associate ourselves only with like-minded people – can be expressed in several possible dimensions, such as those related to socio-demographic dimensions that stratify society ("status" homophily), and cognitive dimensions (e.g., preferences, attitudes, aspirations, values) (“value” homophily), as defined by McPherson et al. (2001)[28]. Status homophily includes ascribed characteristics (e.g., race, ethnicity, sex, or age) and acquired ones (e.g., religion, education, occupation, or behavior patterns). Value homophily includes dimensions such as beliefs and attitudes, traits like intelligence and behavior (emotional), which may report cognitive similarities. In this regard, McPherson et al., (2001) states that value homophily is about "internal states presumed to shape our orientation toward future behavior" (p. 419).

Thus, similar people will establish contact at a higher rate than dissimilar ones[29]. It is assumed that endogenous characteristics strongly affect the creation of ties. In fact, homophily is often studied in the perspective of social ties creation and maintenance in social networks, and so, associated with the empirical measures of assortative mixing. There are three main factors related to why social networks present assortative mixing, which refers to a positive correlation in the personal attributes among individuals who are socially connected. One factor is related to homophily, which justifies why people create foci of shared information and points of view. This socio-demographic and attitudinal information implies that “distance in terms of social characteristics translates into network distance” (McPherson et al., 2001, p. 416). Thus, the stratification of

---

homophily and social influence is not easy to make. Some of the difficulties in distinguishing these phenomena may be related to external factors (difficult to be specified) (Anagnostopoulos et al., 2008).

[28] Following the work of Lazarsfeld & Merton (1954).

[29] Given the principles of homophily, authors have proposed various measures to study dimensions like attitude and background factors (McCroskey et al., 1975, 2006), cultural similarity (Centola, 2007), and they have underlined the relevance of some variables like Educational, Occupational, and Class homophily, comparing less intimate ties to relatively strong ties. For a review see Rivera et al. (2010).

society by similarities and dissimilarities between individuals also means distance to be travelled by a piece of information between two individuals.

A second factor is related to how network structures may influence the formation of social ties. This can happen, for example, through the propinquity mechanism that leads to spatial proximity: this mechanism explains that there is a high likelihood that two people that do not know each other will meet, if they share time with the same third person (common friend). In these circumstances, the physical distance changes the likelihood of tie formation, weak or strong (Hipp & Perrin, 2009).

Finally, there is the property of sociality, which is not related to homophily, but still might influence the formation of ties.

Homophily may lead to cognitive similarities (Arazy et al., 2010) that may affect the perception of the communication. In this regard, some authors (Roger & Bhowmik, 1970; McCroskey et al., 2006[30]) uphold that the communication is more effective when the source and the receiver of information are more similar (homophilous), since the perception of the message is associated with cognitive similarities (Roger & Bhowmik, 1970). Moreover, when the perception of the message is associated with cognitive similarities (Rogers & Bhowmik, 1970), communication becomes even more effective. These similarities, even with limited social interaction, are likely to establish links of trustworthiness that may induce receivers to more comfortably accept a sender's recommendation (Arazy et al., 2010).

Differences in the social context in which people are embedded, also affect communication given the varying levels of attitudinal diversity (Levitan & Visser, 2009). Psychological preference (Kossinets & Watts, 2009) seems to drive one of the reasons why individuals interact favorably with others who are similar – strengthening

[30] Several studies have debated the factors that form the basis of human communication, such as, how the 'person perception' affects the interpersonal communication (McCroskey et al., 2006). The studies report two main factors: interpersonal attraction and homophily (ibid.). McCroskey et al. (2006) have been analysing the reliability of the measures reported from 1975 to 2006, and have concluded that they are valid, while still recommending a second generation of measures. The authors McCroskey et al. (2006) state that the first-generation measures of McCroskey et al. (1975) "reported a multi-dimensional measure of perceived homophily (similarity of source and receiver)" (p.2), but which presented moderate reliability in several studies in which they were used. Thus, the second generation measures (McCorskey et al., 2006) review the measuring instruments for reliability and validity of the homophily scales, namely by analysing thirty years of work of other authors that have used those scales. Furthermore, McCroskey et al. (1975, 2006) suggest that perceived and real homophilious patterns are present regarding the following factors: age, ethnicity, sexual orientation, religious affiliation, education level, income, attitudes, beliefs, values.

ties (Granovetter, 1973), and so emphasizing a cumulative advantage of homophily, if there is a preference for homophilous relationships[31]. By tracking the email of 45,553 students, faculty and staff at a large research university over an academic year, Kossinets & Watts found that Simmel's triadic closure[32] was the predominant influencer on social attachment. However, it should be noted that similarities are often measured through scales of homophily, and psychological attributes in the social network analysis are neglected (Crosier et al., 2012). Diversity is also affected by homophilious behaviour, even in contexts where diversity is explicitly valued and encouraged (Mollica et al., 2003; Ruef et al., 2003; Ingram & Morris, 2007). As reviewed in Rivera et al. (2010), diversity can be found in some heterophilous dynamics, such as boards of directors of large companies, to be representative of dissimilar functional specializations (e.g., law, science, or non-profit), (e.g., Westphal & Milton, 2000, Mizruchi, 2004), teams in science (e.g., Moody, 2004), or even in the formation of task-related ties in organizations (asking for assistance or support from a colleague) (Casciaro & Lobo, 2008).

Therefore, homophily contributes to clustering people that share similar social and cognitive dimensions. Basically, homophily is a natural 'source' of social echo, among similar people. As matter of fact, when people realize the similarities between them, mutual trust is enhanced, but, on the other hand, these people become "naturally" closed in clusters, framed by similar opinions and viewpoints.

Strong ties in Echo Chambers

The strength of the tie is intimately related to homophily, which characterizes socially linked people. This leads to the mechanisms of exposure that are associated with the tie structure in cohesive networks. The exposure between strong ties is associated with the time spend with each other's contacts, i.e., with close friends or co-workers, and the motivation to interact may derive from endogenous effects, i.e., homophily, which are conducive to the formation of ties, or to the sharing of preferences (towards cultural interests, etc.).

[31] Yet, networks which are already highly homophilous and, e.g., exposed to mechanisms of triad closure (Rapoport, 1957), do not easily become more homophilous due to the cumulative effect (Kossinets & Watts, 2009).

[32] This concept is detailed further below and in the next section.

Social structures are depicted by a variety of interactions, either at dyadic and triadic level, such as face-to-face in local groups of neighbors, or even in organizational and categorical social structures. In particular, social structures based on strong ties are characterized by relationships surrounded by strong third-party connections (Reagan & McEvily, 2003). These triads contribute to the principle of triadic closure that strongly affects the formation of ties in social networks, as detailed in the next section. This kind of social structure is typical in cohesive networks and finds its roots in Granovetter (1973). In this regard, the author states that individuals are 'embedded' in a matrix of relations and ties forming cohesive embedded networks. Cohesive networks represent the context in which individual actions are placed and in which individuals tend to interact more frequently and spend more time with each other's contacts.

In general, "strong ties have greater motivation to be of assistance and are typically more easily available [than weak ties]" (Granovetter, 1982, p.113). People linked by such ties are more likely to engage in higher emotional efforts to share knowledge, with others. Conversely, this characteristic of embeddedness may intensify the flow of influence (Bian, 1997) among strong ties. This is the case of the flow of diversified knowledge at dyadic level through strong ties that reinforce the enhancement of individual creativity (Staber, 2004; Sosa, 2011), or even knowledge creation – among university researchers – if strong ties are surrounded by a sparse network of actors (Mcfadeyn et al., 2009).

In all these cases, cohesion is due to the characteristics of strong ties that contribute positively to the flow of specific resources, but, on the other hand, which constrain access to new information and diversity of points of view. Similarly to what happens in offline social structures, people in online social networks keep the same quantity and diversity of close friends in their core networks, with whom they communicate most frequently and from whom they receive the majority of information. Thus, the dynamics of communication do not mean an increase of new close friends (Wang & Wellman, 2010). In fact, it has been reported by scholars that online social networks encourage communication with existing offline connections, instead of being a "trigger" to initiate new contacts online (Ellison et al., 2007; Subrahmanyam et al., 2008). Therefore, this dynamic of communication that appears to maintain the same group of strong ties, intensifies the echo chamber effect between these individuals, keeping those that do not share the same viewpoints apart.

Research in this vein has shown that strong ties are an important determinant of attention on social networking websites, such as Facebook (Messing & Westwood, 2013), and that Facebook users usually browse profiles of people with whom they have an offline connection more than the profiles of complete strangers (Lampe et al., 2006). Furthermore, scholars report that acquaintances in networks have a determinant role as vehicles of contagion, due to their abundance (Bakshy & Rosenn, 2012), but also to give access to novelty (Granovetter, 1973). However, these ties are not the ones that characterize the current use of social data for Web personalization.

As matter of fact, when users browse their social network, they access content posted by friends, acquaintances and through them, from friends' friends (third-party connections). This means that a large network of relationships is established between them all, which include people (senders and receivers) and the exchanged contents.

As a result of this, the correlation between social network structure and users' attributes emerges, in which people with similarities exchange and access similar kinds of content. The structure formed by such ties, finds correspondence in users' attributes as shown by Mislove et al. (2010). The authors posit that using given attributes from a fraction of users in an online social network, it is possible to infer the attributes of the remaining users. Network communities form around users who share certain attributes. Thus, given the shared cognitive similarities in such groups of people, the information spins in closed circuits. Therefore, this seems to interfere with the effect of Echo Chamber, by emphasizing it.

Network mechanisms in the formation of ties that feed the echo chamber

The two most studied determining factors of the formation of ties are triadic closure[33] and selective mixing. Selective mixing is related to the tendency of tie creation based on people's attributes (e.g., language, homopily dimensions) (Goodreau et al., 2009). Both factors are strongly supported or reinforced by homophily factors (Schafer, 2011). Two known factors determining tie formation are the structural proximity (e.g., friendship circles, shared foci) and physical distance (Marmaros & Sacerdote, 2006).

[33] A triad can be described as a set of three people that tend to close through a third person due to propinquity or cognitive processes.

This section centers attention on the consequences of triadic closure given its importance in understanding the effect of echo chamber.

More than 100 years have passed since Simmel (1908) showed the importance of triadic clustering in social interaction, and since then many scholars have addressed this issue in social network analysis (e.g., Granovetter, 1973; Wasserman, 1974). The basic idea of triadic closure is supported by the Balance Theory of Heider (1958), which posits that two people may appreciate each other mutually by way of their mutual agreement on a third person. This third person is a common friend, or someone with whom the two others spend time together. Supported by this theory, Granovetter (1973) explained that if an individual B is a friend of individual A, and A of C, then, there is a high probability that B and C are friends (or become friends). These interactions among friends and people with similar interests or behavior have been studied by sociologists namely in the context of tie formation. All those processes are often characterized by the interplay among homophily dimensions and tie strength, it being commonly accepted that individuals seek or join groups that are close to their prior beliefs and biases, as seen above.

Two balance mechanisms contribute strongly to tie formation: reciprocity – which includes the desire to reciprocate the friendship (Granovetter, 1973), and transitivity (e.g., ethnic homogeneity on online social networks) – that describes the tendency for friends-of-friends to become friends (Goodreau et al., 2009). Both mechanisms contribute to measuring similarities among people (their homophily) and describe a certain closure among similar people (friends) that contributes to tie creation.

The representation of those individual characteristics are extensively reported in literature, such as in the context of adoption of behavior (Zhou et al., 2009), contagion (Aral et al. 2009), or creation of ties by people's similarity or dissimilarity (Rivera et al., 2010). Furthermore, scholars (e.g., Golder & Yardi, 2010; Leskovec et al., 2008) have been proving that friendship connections among users in their online social networks are mostly based on a triadic closure principle, i.e., people mainly form connections with, or through, close friends (strong ties).

Leskovec et al. (2008) study the triadic closure mechanism within four online platforms[34] to contend that triadic closure justifies the most links between people in large online social networks. They found that most of the new edges (connections) are extended to very short distances, typically close triangles, through which is possible to present a predictive model of network evolution that captures the triadic effect.

In the same vein, Golder & Yardi (2010) show that transitivity and mutuality emphasize the effect of triadic closure among Twitter users, and Gilbert (2012) reports that the formation of tie strengths manifests itself in similar ways on Twitter and Facebook, which shows the triadic closure principle in Facebook as well. The author's findings can be generalized across media, revealing too that some important properties of online relationships do not change due to implementation details on SNSs, e.g., changes on design and functionalities, like those observed on Facebook since 2008.

Thus, scholars stress the role of strong ties in tie creation, as well as how people rely on them given the shared trust, which is confirmed and reinforced by network mechanisms and endogenous factors. In fact, all this is explained by the browsing activity and communication level between such ties.

4.4 Social Echo Chamber Effect on Web Personalization

The reported findings on tie creation in triadic closure stress the argument for the relationship between the effect of echo chamber and the use of social data from strong ties and homophilous people – through which ties are set or are in a state of being established by triadic closure. Consequently, if the data from social networks are based on profile similarities and people socially connected by strong ties to improve the performance of Web personalization, the final results are the known effects of echo chamber.

A related problem associated to personalized recommendations has already been studied under the known concept of Filter Bubble (Graells-Garrido et al., 2013), as detailed in the previous chapter. In this regard, this authors test a way to take advantage of partial homophily between people with opposite views on sensitive issues. After determining the regularity of keywords between people with different viewpoints, the

[34] 1) The photo-sharing website: "FLICKR" (flickr.com); 2) The collaborative bookmark tagging website: "DELICIOUS" (del.icio.us, a); 3) "YAHOO! ANSWERS" (answers.yahoo.com); 4) The professional contacts website: "LINKEDIN" (linkedin.com).

authors created a data portrait of each user, and then recommended information based on similarities between their word clouds, especially when they differed in their views on the topic studied (abortion on Chile). The results show that people can be more open than expected to ideas that oppose their own, which is a relevant approach to disrupt the Filter Bubble.

Another known problem that affects users' satisfaction is called the Portfolio Effect. This term was coined specifically for recommender systems referring to recommended items that were already familiar to users (Burke, 2002). One example of this effect can be found in news recommendation. Here the recommendation lists often contain identical or nearly identical news messages only.

Another example appears in recommendation engines like *Amazon.com*. In the case of *Amazon*, the effect is found in costumers that have purchased several books from the same author, which may bias the recommender system. From that, the users may receive recommendation lists where all top-5 entries are books by that author (Ziegler et al., 2005).

A similar constraint happens with content-based recommender systems, especially with respect to music, where songs of the same artist are recommended. In this regard, Seyerlehner (2010) proposes a solution based on the use of a portfolio filter aiming to ensure that there is only one song per artist in each recommendation list and then force the content-based recommender systems to increase the diversity of the recommendations.

Although there is a diversity of terms, methods and solutions to solve the different problems identified in web-based systems, notably in recommender systems, this dissertation indicates a new problem and suggests a solution to solve it. As seen, this problem derives from the use of social data to improve personalized recommendations through social-based recommender systems. However, as a result of network properties, particularly given the interaction results of individuals at emotional and cognitive level, the current solutions applied in social-based recommender systems end up creating an Echo Chamber Effect that traps people inside their – usual – social bubble of information.

The Social Echo Chamber Effect refers to the fact that users are trapped inside their own social of information, received and known. This loop of information keeps users attached to the same viewpoints and away from the access to novelty. As seen, this may occur when the target-user receives a recommendation based on information from similar users or close friends (strong ties), or simply by mirroring their past online activities. Research in social psychology has argued that our identity is shaped by the media we consume. In short, identification processes are powerful engines in engaging the users during reception (e.g., Entman, 1989; Cohen, 2001). So, we can easily fall into feedback loops that we are not aware.

Scholars do not mention what the individuals' options on content selection would be if the individuals could opt from a sender weakly or strongly tied to them. It is also not reported what a possible benefit at cognitive level could be, notably considering the meaning construction of the receiver if the selection of content were free of algorithms options (which may be biased and serve unasked queries by the users). This means, free of information tailored by past options, with the aid of technology, about people's tastes, views, and prejudices.

However, there are reports on the role of weak ties in contagion processes and access to novelty, which reinforces the view put forward in this dissertation. These and other arguments are tested in the empirical works presented in the sixth, seventh and eighth chapters. I conjecture that the lack of studies in Social Network Analysis in this area might be one the reasons why the Echo Chamber Effect – related to the use of social data – has been absent from literature so far. In the following chapter an overview on social networks is presented in order to introduce the empirical chapters.

CHAPTER 5

SOCIAL NETWORKS OVERVIEW

5.1 Overview

Social Network Analysis (SNA) was initially formalized within the frameworks of graph theory and network theory. The perspective of SNA that includes both method and theory claims that there is no meaning in studying any single relationship in isolation from the network of which it is part. In fact, the dyad formed by the relationship between two individuals is the main element of a network, but its existence is itself conditioned by the network. The SNA methods and underlying theories on social science frame the empirical work of this dissertation. Additionally, some psychology theories are applied that challenge the current studies on SNA, notably on bridging ties.

The growth of computing power and the current trend of social media that mirror the engrained desire of humans in connecting on large scales (Crosier et al., 2012), has opened new possibilities for accessing massive amounts of data and going further in SNA studies. In this venture, several types of network-oriented mathematical software have been developed to assist the work on SNA. Some examples include the UCINET, NodeXL, statnet in R, Pajek, or EgoNet, just to quote a few of them. Furthermore, the use of both the theory and methodology networks has already crossed the boundaries of social science and reached multiple fields (for a revision see: Kadushin, 2012). Its application has become interdisciplinary and has motivated the adoption of new methods in numerous scientific areas, like psychology (Vachon, 1982; Kalish & Robins, 2006; Brass, 2011), human behavior (Li & Chen, 2014), communication (Oberg & Walgenbach, 2008; Vladuțescu, 2012), social media & emotions (Kivran-Swaine & Naaman, 2011; Lin & Qiu, 2012; Tadic et al., 2013), biology (Fowler et al., 2009), health (Cornwell, 2009; Haas et al., 2010), organization science (Shipilov, 2009; Ahuja et al., 2012), economics (Jackson, 2010; Ozsoylev & Walden, 2011), or even behavioral ecology (Sih et al., 2009).

Networks of relationships come in many shapes and sizes, which complicate the task of finding a single way of representing them encompassing all applications. Yet, there are some common network representations that help to accomplish this purpose. Next, and in order to introduce the social networks concepts applied in this dissertation, notably in the three empirical studies presented in the following chapters, an overview on social networks and the fundamental theories behind tie strength and central nodes is presented.

This chapter is organized into six sections. The first, *Networks of relationships* presents a short overview of some elements that denote a social network in order to introduce the other sections of this chapter. The second, *The Strength of Weak Ties*, presents the fundamentals of the theory, some approaches of other authors and the trade-off between weak and strong ties. The third, *Network Bridges,* introduces the concept of network bridges. Two other topics are developed here: *Bridging factors through weak ties* and *Bridging factors through non-redundant structural holes*. The fourth section, *Central Nodes: centrality and bridging measures,* outlines the differences between the network measures defined by the concepts of centrality and bridging. The fifth section, *Size and tie diversity,* overviews the concept of size beyond the notion of number of network members. It presents the relationship between size and weak and strong ties, and how it contributes to diversity and determining the value that a user can derive from being a member of a network. The last section, entitled *Psychological attributes in social networks,* presents some studies and reflections on the merging of both fields of research.

5.2 Networks of relationships

A network is a set of relationships (Kadushin, 2004), while a social network can be defined as a "finite set or sets of actors and the relation or relations defined in them" (Wasserman & Faust, 1994, p.20). Social networks operate on many levels, from individuals and families up to the level of nations. They play a critical role in determining the flow of information through central nodes, and the degree to which individuals succeed in achieving their goals. They also have an influential role in the way problems are solved, or in how organizations are run. This social structure made by nodes is viewed in general as individuals or organizations tied by one or more specific types of interdependency that includes values, visions, ideas, financial exchange, friendship, kinship, dislike, conflict or trade (Bulte & Wuyts, 2007).

“Communication networks are the patterns of contact that are created by the flow of messages among communication through time and space” (Monge & Contractor, 2003, p.3). The concept of message includes here diverse symbolic forms (e.g., data, information, knowledge, images, or symbols) that flow between points in a network or that can be co-created by network members. These networks take many forms in contemporary organizations, which contain personal contact networks, or flows of information within and between groups, to name a few (Monge & Contactor, 2000).

Social network data can be applied in the construction of both personal (ego) and whole networks, wherein they both share several measures. However, each one of the networks has a unique set of structural metrics given the specificity of each social system being modeled.

In personal networks random sampling methods are adopted to define the boundary of work and to make the work of data collection more feasible. However, the dynamics related to personal networks are complex given the multitude of interactions associated to the relational level. Here it is the owner of the network, the ego, that generates the list of members (‘alters’) of their own social network, which change in size, composition, structure, and stability. The ties can be created, grow or decrease in strength or change their contents, but also disappear in a smooth way or end abruptly. One of the challenges related with this kind of networks is the fact that only one person informs about the network, which can make hard to predict any change that occurs in the network.

A whole (‘sociocentric’) network is considered an entire population of individuals bound by a concrete definition (e.g., students who attend a school). The increase of members is accompanied here by the number of possible connections between them and then by the size of the group. Here, measures of structural holes (bridging) can be represented by betweenness centrality measures (Ferris and Traeadway, 2012). Some scholars associate these measures with power in organizations (Brass, 1984), while the measures of structural holes in personal networks have shown robustness in predicting performance outcomes (Burt, 2007).

One of the benefits of analyzing social networks is that they can help researchers to understand and evaluate how structural and relational properties and individual attributes intervene, notably on perception of information, access to novelty, or even on sharing information and knowledge. For example, with respect to the access to novelty, a network rich in acquaintance connections, i.e., weak ties (Granovetter, 1973) and structural holes (Burt, 1992) can be considered an indication of bridging factors that represent network central nodes. These analyses are intimately related to the theory of '*The Strength of Weak Ties*', as is shown next.

5.3 The Strength of Weak Ties

Some of the principles that describe relationships among individuals in a social network are found in 'The Strength of Weak Ties' (Granovetter, 1973). The author states that within social circles there are individuals who establish ties with members of outside networks. These ties are called 'weak ties' because they are built by distant individuals who can still give access to each other's resources, and end up becoming strong ties instead.

Granovetter's work proposes a measure of tie strength, which has the underlying principle that personal ties have an important association with reciprocity, in the sense that "the strength of a tie is a combination (probably linear) of the amount of time, the emotional intensity, the intimacy (mutual confiding) and the reciprocal services which characterize the tie" (p. 1361). In this work, the tie strength measure was calculated by asking those who found a new job through contacts how often they saw the contact around the time that he passed on job information to them. A scale of three items was conceived: strong tie = at least twice a week; medium tie = less than twice a week but more than once a year; and weak tie = once a year or less.

Granovetter' survey was centered on understanding the relevance of tie strength in finding jobs. Granovetter interviewed people (n = 54) who had found their most recent job through a social contact. This work, which has become one of the most influential research projects in the field of social networks, found that people got a job more easily through weak ties contacts (27.6 percent) than strong ones (16.7 percent). Medium ties represented 55.7 percent of the contacts. The findings contrasted the differences between weak and strong ties and their role in society.

The results may have a different trend when considering the urgency for a job. In this circumstance, Granovetter (1983, p. 211) says that people "in urgent need of a job turned to strong ties because they were more easily called on and willing to help, however limited the information they could provide".

Two authors (Murray et al., 1981, Bian, 1997) stated a view contrary to Granovetter's theory applied to job finding. Nevertheless, in both, it seems that these findings are different from Granovetter due to two main reasons: a) particularities of the sociocultural and temporal context; b) understanding from the sampling obtained.

In response to Murray et al. (1981), Granovetter (1983) argues that 80 percent of the data of these authors are focused on first academic jobs. Thus, because the new PhDs have, in general, few useful contacts in their subject, they need to rely on mentors and dissertation advisers, with whom they have a close relationship – at least at academic work level. By observing the percentages and number of PhDs for no first job and estimating these figures after disaggregation by career stage, Granovetter posits that their reliance on strong ties should decline, confirming Granovetter's (1973) theory.

The other author, Bian (1997), argues that strong bridging ties are also efficient when it is the influence that flows through personal networks instead of information. In such cases jobs can be channeled through strong ties more easily than through weak ties. The author studied a particular socio-cultural context, where personal networks are used to gain influence from job-assigning authorities, rather than to gather employment information. The distinction between information and influence may disentangle some controversies about the relative efficacy of strong and weak network ties in the context of job searches. Nevertheless, the author concludes that, despite the strong-tie bridges observed in his work that challenges the strength of weak tie hypothesis, so immersed in Granovetter's work about job searches, this does not totally disqualify it. Here, it is relevant to note the socio-cultural context of China, where the study was undertaken. As mentioned by the author, in China "personal networks are used to influence authorities, who in turn assign jobs as favors to their contacts, which is a type of unauthorized activity facilitated by strong ties characterized by trust and obligation" (p. 366).

Since Granovetter's theory, scholars have debated several proposals to measure tie strength in a social network. Often, the concern in describing the 'strength' of ties is in identifying how people are close to or distant from each other. Another important topic of debate has been the possible patterns related to structural positions of the individuals in social networks and the attempt to predict actors' roles.

Lin et al. (1981) state that strong ties can be defined by social distance – embodied by factors such as socioeconomic status, education level, political affiliation, race and gender.

In turn, Marsden (1987) highlights that emotional closeness is what can best reflect tie strength, while Wellman et al. (1990) characterize strong ties as those that offer emotional support, such as offering advice on family problems.

Krackhardt (1992) argues that strong ties can be depicted by interaction, affection and time. This author also explores how people's behavior in processes of information sharing can generate trust.

Burt (1995) claims that tie strength characterize the network topology and the informal social circles. While Gilbert & Karahalios (2009), Ficher (2010) and particularly Krackhardt (1992), claim that strong ties are defined by seven main dimensions: intensity, intimacy, duration, reciprocal services, structural, emotional support, and social distance.

Last, but not least, Petrosky (2011) states that "strength" can be conceptualized as being consisting of two dimensions: intensity (frequency of contact) and valence (the affective, supportive and cooperative character of the tie).

From the extensive literature published related to tie strength results, both weak and strong ties are important channels through which users extract benefit from their networks. It is known that weak ties show great utility in searching for information and that their value derives from locating what needs to be exchanged.

On the other hand, strong ties are useful for exchanging effective or tacit information and on making exchanges (Hansen, 1999; Granovetter, 1985). In this sense, there is a trade-off between the opportunities to access new information through social distant ties (weak ties) and the micro integration that allows the regular transmissions within groups (Friedkin, 1980) into which strong ties are usually immersed.

Strong ties are then not less relevant in a network. They are known to have greater motivation to be of assistance (Granovetter, 1982) and influential in determining the outcome of a union election (Krackhardt, 1992). Strong embedded ties have good problem-solving capabilities, particularly when compared with other nearby connections, afford higher levels of trust, and are good conveyers of information (Uzzi, 1997). Although, strong ties can be a more trusted source of advice and influence in uncertain or conflicting situations, they require more time and effort to keep, which may originate stronger obligations to reciprocate than weak ties (Ferris & Treadway, 2012), but also keeps each one more aware of the others' viewpoints and information.

Social network users usually compete for attention and rely on each other to spread their information and contents. Such contents gain importance depending not only on their quality, but also on their size and spreading process.

In this scope, Shi et al. (2013) show that Twitter followers who are weakly tied to content senders are more likely to retweet than strongly tied followers. Other authors explain popularity (individuals with large numbers of friends and high volumes of communication) as being inversely related to picking and sharing contents on Twitter (retweets) – to the extent that, when an individual becomes more popular, their rate of retweeting goes down, particularly when they are followed by a large number of people (Harrigan et al., 2012). These results drew our attention because of the relationship between people's behavior and the strength of the tie, whose interdependence can be associated to the conjecture that a pair of strongly tied people shares a larger overlap in their friendship circles – reducing novelty – than a pair of weakly tied people (Granovetter, 1973).

While some authors report that reciprocal ties, or ties with common third parties that are common in community structures, substantially increase social contagion in social networks such as Twitter – users are more likely to disseminate redundant information (retweeting "old news") (Harrigan et al., 2012) – others show that most contagion occurs along weak ties, given their abundance in social networks (Bakshy & Rosenn, 2012). Moreover, weak ties are the best channels for gaining access to novel content that people would otherwise not find (Bakshy & Rosenn, 2012) and conduct useful information in computer mediated communication (Constant et al., 1996).

Finally, the strength of the tie between sender and receiver is also reported as being a strong determinant of attention to traditional media items on social networking websites, e.g., Facebook (Messing & Westwood, 2013). The authors' findings suggest that social influence serves to privilege information shared by socially close friends at the expense of heterogeneous contacts, being that a powerful force driving news consumption. This study attempts to do a direct causal examination of how the strength of the tie between sender and receiver drives attention, independent of common interests or other sources of similarity/homophily.

Often, the strength of the tie is characterized in published literature by seven dimensions: intensity, intimacy, duration, reciprocal services, structural, emotional support and social distance (Ficher, 2010), whereas weak ties are commonly associated, in a simplified way, to the infrequency of contacts between the individuals and capacity to traverse greater social distances establishing local bridges (Granovetter, 1973).

5.4 Network Bridges

Weak ties are characterized by the infrequency of contacts between the individuals and their capacity to channel ideas, influence and novelty by traversing greater social distances. Moreover, individuals connected by ties who do not share ties with other people of the same network are local bridges (Granovetter, 1973). Though, not all weak ties are bridges as noticed by the author. Hence, when weak ties are "bridges" they become sources of information that can bring new perspectives and create new insights, which strong ties cannot. This happens because people often share their opinions and perspectives within the social circle linked by strong ties (e.g. family and close friends); yet, since the strong ties will already be familiar with their ideas, it reduces the possibility of accessing different viewpoints in the process of information sharing.

Granovetter states that a strong tie can be a bridge, but only if neither party has any other strong ties. Furthermore, if the individuals of a network form triangles through their connections, formed by transitivity, then it is not possible to establish local bridges between them. The author claims that the transitivity mechanism can be regarded as a “function of the strength of ties, rather than a general feature of social structure” (p. 1377).

Connections established at triadic-level forming triangles shaped by transitivity, reinforce the strength of the ties and their proximity due to similarities between the individuals in these closed circles. This reduces, for those ties, the ability of acting as a bridge (hinge) with other social circles. This is a common circumstance in endogenous and structural conditions that contribute to tie formation in social networks, i.e., triadic closure and selective mixing, which are strongly supported or reinforced by homophily factors (Schafer, 2011). Consequently, ties that involve little time, effort, and emotion (requiring little pressure to organize activities with others) to stay connected, are most likely to remain bridging (Feld, 1981). This premise is contrary to what happens between strong ties.

Bridging factors through weak ties

In an amplified re-edition of his 1973 work, Granovetter (1983) says that scientific discoveries are more able to flow through weak ties than through strong ones. Similar conclusions are reported by Lin et al. (1978) that re-created Milgram's experiment of the "The small world problem"[35]. This experiment was based on a request to their participants to deliver a booklet to some unknown person in a distant place. The authors confirmed that weak ties were more efficient in helping the booklet to reach the destination.

In a different vein of investigation, Ruef (2002) confirmed the relevance of weak ties for creative entrepreneurs to achieve non-redundant information that will contribute to innovation. Those weak ties – of acquaintanceship, of colleagues who are not friends – may act as bridges between non-connected social circles.

Granovetter's bridging concept was also discussed and found to be beneficial for the overlap of several sub-networks with many others affecting the motivation of employees in their work places (Blau, 1980). By studying the integration of staff in a children's psychiatric hospital in New York City, the author reports that good integration of employees (contrary to comparable institutions, there is not a high staff turnover, neither a low morale) can only be understood by considering the role of an extensive network of weak ties. She found a correlation between the network of weak ties and low staff turnover with high morale. If instead of weak ties there were strong ties, and given the sub-networks of many different foci (i.e., hospital departments),

---

[35] In approaching the work of Milgram (1967) – "The Small-World Problem" – it becomes clear that through an average of five circles of acquaintances apart is possible to reach anyone on planet (i.e., six degrees of separation). This work confronted two different philosophical views of the small-world problem. One posits that two people can be linked through acquaintances, and that the number of such intermediate links is relatively small. The other holds that there are unbridgeable gaps between various groups. The author concludes that "social communication is sometimes restricted less by physical distance than by social distance" (p.66). Because this work was deeply embedded in the cultural context of the mid-century United States, Milgram raised the question about possible differences in the results if the experiment would take place in a different society with more sustained kinship relations. The answer was given by Lin et al. (1978) and reinforced by several other investigations, e.g., Watts (2004), which posits that many real-world networks, as social networks, could be small-world networks.

these sub-networks would tend to close in on themselves, and then they would develop into cliques, as highlighted by Granovetter (1983) discussing Blau's work. Such cliques can create closed circles of information flows and of personal interactions, which could reduce productivity and employees' motivation. Blau's work highlights the value of weak ties in social interaction and an important relationship between psychological health (high moral) and network structure.

Bridges formed by weak ties also have a positive impact on individual creativity (Perry-Smith, 2006) and on keeping a low redundancy in the flow of information (McEvily et al., 1999; Hansen, 1999; Ruef, 2002). Scholars also report that weak ties are more prevalent in structurally diverse networks, being determinant for the diffusion and propagation of novel information (Bakshy & Rosenn, 2012).

As has been seen, there are abundant studies testing the hypotheses put forth by Granovetter, in particular on the role of weak ties as a bridging factor. In this scope, many interesting questions have been answered by scholars about the relative use of weak ties, but some have still not been fully answered. For instance, in this dissertation the question of the role of weak ties on recommending surprising information is raised. Is the importance of the weak ties only centered on their bridging behavior or ability to diffuse information, or do they embody other features like an "emotional opportunity", i.e., surprise, that can be expressed in a regular structural distance, or cognitive distance to other people? Should this 'distance' be considered only from a structural perspective, or cognitive (personal attributes) or both? This subject is debated carefully in the next three chapters. In particular, cognitive distance is discussed in chapter eight.

Meanwhile, as seen, considerable literature has been published about the tie strength argument claiming that weak ties can provide people with better access to information and resources beyond those available in their own social circle (Granovetter, 1983). Since then, this has been the most common approach to expounding the benefits of bridging ties, although it is not the only approach that highlights the benefits of network bridges.

Bridging factors through non-redundant structural holes

A second network theory on bridging ties was developed by Burt (1992), which introduces the concept of structural holes. It is argued that the individual who establishes a bridge between two acquaintances not connected with each other provides superior access to information and greater opportunities to exercise control. Thus, the structural holes through which new information flows, also lead to inequality between network members and power opportunities. Individuals with different attributes and organizations of different kinds may not be affected equally by these holes.

A structural hole is a void in a social structure. In terms of social networks this refers to an absence of connections between individuals and each one of them being the access to different groups. This does not mean, though, that these individuals and corresponding groups are unaware of each other, but rather that the lack of links between them leads to a non-redundancy in the exchange of information. Thus, as Burt points out, receivers of less redundant information through individuals that span structural holes are better informed about opportunities and hold a broader range of options to access diverse individuals whenever it is worthwhile.

Contrary to Granovetter (1973), Burt's theory introduces a measure of bridging that is a function of the redundancy of contacts between individuals that span structural holes. This measure calculates the spanning function by "constraint" (p. 55). Constraint is the degree of redundancy of the contacts of an individual. Such contacts are redundant to the extent that they lead to the same people, and so provide the same information benefits. This measure is positively related to the formation of structural holes, where a high value of constraint means more structural holes (Susskind et al., 1998). This measure of bridging can also be evaluated through triadic-level measurements which can become advantageous when establishing comparisons across networks (Kalish & Robins, 2006).

Burt (1992) asserts that individuals acting as brokers have control advantages over the information flow and, as brokers, are the third person, in the established connection, and strengthen their position by benefitting from the information shared between receivers and the originator of the information. In this regard, and considering a multidisciplinary viewpoint (i.e., health and social networks), Cornwell (2009) advances that the advantages of being a broker in one's own network depends on the individual’s mental condition, because bridging actions use the ability to recall or identify the structural holes of the individual’s network. Many other scholars have been debating the role of structural holes in several fields of application, including the discovery of new information (McEvily et al., 1999), the access to novelty (Gilsing et al., 2008) and its delivery (Aral & Alstyne, 2011), or even how social network structures may influence people’s outcomes, such as creativity (Burt, 2004; Uzzi & Spiro, 2005; Fleming et al., 2007; Sosa, 2011).

In summary, as we have seen, some studies report similar outputs to both bridging factors (e.g. creativity), but there has been less research into finding out whether or not both factors are equally related to the perception of novelty by receivers. This will be discussed in the next chapter.

5.5 Central Nodes: centrality and bridging measures

Two structural positions determining the flow of information in social networks have been described in literature. These two types of central nodes can be measured by: a) centrality (Freeman, 1979) and, b) bridging factors (Burt, 1992).

a) Centrality is defined as the extent to which individuals are connected to others in a direct or indirect way in a network (Freeman, 1979) and posits that individuals who have more ties to others may be in an advantageous position to make many others aware of their views, to hold direct access to resources and show independence from others (Brass & Burkhardt, 1992). These central positions are considered to be preferential given that they represent control or better access to resources (Paruchuri, 2010). Thus, individuals in such central positions contribute to the interconnectedness of the overall network (Rogers & Kincaid, 1981), holding a certain level of power (Brass, 1985; Krackhardt, 1990) given their easy and direct access to any resources that might flow through the network (whether or not dependent on any particular individual). This general view of network centrality suggests that the benefit of a central position depends on the interdependence maintained with the adjacent nodes. Two common ways to measure this are by the number of relationships or the size of one's network. Both are referred to as degree centrality (Ferris & Traeadway, 2012).

The metrics most often studied to characterize centrality were introduced by Freeman (1979). They include degree, betweenness and closeness centrality[36]. Degree is a local measurement, undertaken at dyadic level and focused on the level of interaction, e.g., of the communication activities. It can be calculated by counting the number of links for each node. Often, it is interpreted as a grade of popularity, prestige, or influence (Knoke & Burt, 1983), and it is argued that the influence exercised must be related to a higher degree and clustering coefficient value – the followers have to be linked between each other (Kanovy & Yaari, 2011). Others report that it can be indicative of the avoidance of relying on mediating nodes for indirect access to resources or even other direct interactions such as coalitions (Brass & Burkhardt, 1992).

Betweenness and closeness are global measures and are calculated using information from the entire network. Betweenness centrality is frequently observed from the broker's standpoint, which is positioned on informational paths facilitating the flow of information and connections between individuals (Mori et al., 2005; Kratzer & Lettl, 2008). Formally, this measure refers to the probability that a 'communication' between two individuals takes a particular path. It is assumed that the connections have equal weight, i.e. each tie has a weight of 1, and communications will flow along the shortest paths. These paths minimize the number of intermediary nodes and its length is defined as the minimum number of ties linking the two nodes, either directly or indirectly. Thus, a node that holds a high degree of betweenness centrality refers to the number of shortest paths that it facilitates and supposes that a communication that takes place in this way follows one of the geodesics (Wasserman & Faust, 1994).

Closeness centrality measures the mean geodesic distance (the shortest path) between an actor and all other actors in the network (Wasserman & Faust, 1994). Similar to closeness, betweenness is also concerned with shortest paths, but it looks at the fraction of shortest paths that must pass through the ego to be connected. Closeness expresses the ability to avoid being influenced by others. A low value in closeness means shorter distances from others and can be regarded as power to influence (Holme & Ghoshal, 2008). In this sense, shorter distances could also mean faster access to novelty spreading in a network; however, because closeness only

[36] Other alternative measures, such as Bonacich and eigenvector, also take into account the centrality of alters.

ponders connected graphs, the flow through bridges is not considered. Closeness is not analyzed in the scope of this study.

Despite the different interpretations of centrality measures (e.g., Freeman, 1979; Bonacich, 1987, 1991; Brin & Page, 1998), all scholars agree that centrality is a node-level construct (Borgatti & Everett, 2006), whose measurements must fit the type of “thing” that flows in the network (Borgatti, 2005), e.g., virus, or information in social networks or through email exchanges (Wu et al., 2004) and that provides both a visual and a mathematical analysis of human-influenced relationships (Abbasi & Altmann, 2010).

Centrality means a balance between the peripheral position and the central position in a network that mediates a small number of direct contacts with the core of the network (border position) with a high number of direct contacts (core position) (Kratzer & Lettl, 2008). Conversely, nodes rated with high values of degree and betweenness tend to be located in the network’s core (Hwang et al., 2008). In this sense, from a cognitive perspective, individuals in central nodes have better knowledge of the network than those in peripheral locations (Krackhardt, 1990). These individuals, for example, are better informed about others’ knowledge and network to approach or avoid forming coalitions (Ferris & Traeadway, 2012).

b) Like centrality (Freeman, 1979), structural bridging is also a central node. Like degree, bridging is also measured at local level. As proposed by Burt (1992) it can be measured by “constraint”, which is the degree of a person's links (ego network) to people not connected to another. In order to introduce a measure for bridging using complete network data and independent of degree, Valente & Fujimoto (2010) propose a new approach, justified by the importance of bridging behavior to interpretation of network structure and diffusion. The authors state that, as a global measure, betweenness does well at finding bridges as long as the links between disparate groups come from the center of the network. However, when they come from the periphery the existing measures of centrality are not accurate enough to identify such critical connectors, and constraint cannot do so from a global measure perspective. In this sense, the authors propose an alternative bridging measure that calculates the difference in cohesion (inverse of the average path length distances).

5.6 Size and ties diversity

The importance of network size in social networks is intimately related to communication and may reflect aspects of personality, such as larger orientation for socialization, in spite of the fact that larger networks may be harder to maintain (Tillema et al., 2010). This means that size, which is closely associated with the rise of the number of weak ties in personal networks (Hampton et al., 2010), might change people's communication habits. Similarly, the communication level with a larger number of strong ties increases, but the number of new close friends does not (Wang & Wellman, 2010). Scholars posit that people use social networking sites (SNSs), like Facebook, primarily to keep or reinforce existing offline contacts (Lewis et al., 2008), and that people have the same close friends either online or offline (Hampton et al., 2010). Is this beneficial for information access and diversity?

Size is not adverse to the regularity of contact among strong ties (Tillema et al., 2010). People keep the same quantity and diversity of close friends in their core networks with whom they communicate most frequently and from whom they receive the majority of information (e.g. posts on Facebook) (Lewis et al., 2008). Hence, people are increasing the sharing of knowledge among their close friends, and at the same time, they are more exposed to the information from people with whom they are weakly tied. As argued in previous studies, among other factors (bureaucratization, population density, and the spread of market mechanisms), the development of the communications system has increased the number of weak ties, a fact that has been reinforced with the success of social media services (Pool, 1980).

Individuals that are more exposed to larger networks are more exposed to a larger number of weak ties and so less likely to be redundant and more likely to be information rich. This corroborates with Levitan & Visser (2009), who studied how college students would react to different social contexts containing varying levels of attitudinal diversity. It was found that social resistance to attitude change is inversely proportional to the attitudinally diverse social networks. It means that greater attitude stability will imply more attitudinally congruent networks (Levatan & Visser, 2009). By analyzing the social networks of like-minded connections these authors conclude that "the social context in which people are embedded has important implications for the durability of their attitudes" (p. 1058).

Size may imply more diversity if the increase of contacts is based on weak ties, since people exposed to a greater number of social contexts through the information brought by weak ties will be more available to change their attitudes. Consequently, a network structure rich on non-redundant structural holes may leverage the number of contacts (Afuah, 2013). This signifies more exposure and more access to valuable information.

Size is then beneficial for diversity and so for novelty access. It fosters the diversity of the network and affects the people with whom there is a connection. Through weak ties and structural holes, the probability of an individual reaching different people is higher. This means that, by accessing a greater number of different status groups, the diversity of information (and social support) to which an individual will have access will also be increased (Burt, 1983).

Size also gives a measure of social integration being represented by the number of alters (members of a given social network) with whom an individual has a specified social relationship (Marsden, 1987). Hence, size is about the number of network members, but this fact alone, may not be enough, in particular, to determine the value that a user can derive from being a member of a network. Therefore, a focus on network size, for example, without considering the number and nature of ties within the network, can distort reality (Afuah, 2013). Further, though early research focused on the phenomenon of network effects, centered primarily on the role of network size, more recent works claim that other factors, such as structure, need to be considered. Structural factors (centrality of members, structural holes, network ties) and conduct factors (opportunistic behavior, reputation signaling, and perceptions of trust), shape network value, which raises its importance as a driver of strategic action during the life of a network (Afuah, 2013). Additionally, this author posits that an individual that has a central location in the network or bridges structural holes can bring more value to the network. As a result, these members (its structural position) will be more relevant than an undifferentiated member contributing only to the increasing of the network size.

5.7 Psychological attributes in social networks

As seen above, the study of social relationships provides rich data and knowledge to extend the understanding of the matrix that embeds people’s interaction. Is this enough to comprehend such interactions?

Research into social networks is still growing interest in many fields, but not so much in psychology, notably, when compared to sociology, anthropology or even epidemiology, for example (Totterdell et al., 2010). However, two significant areas of work have been receiving important contributions from psychologists investigating community and organizational fields. The former includes work on the interrelation between physical proximity and similarities, beliefs and attitudes, amount of interaction, and affective ties. The latter, includes work on interaction between personality and network position (Borgatti & Foster, 2003).

Some studies in psychology have centered their attention on the relationship between personality traits and network factors. In this regard, the Big Five model is considered to be the best framework to study personality (John et al., 2008). Inclusively, this framework was tested on online social networks showing that people do not use an idealized virtual identity to interface with others through these communication platforms, which suggests that they might be an efficient medium for expressing and communicating real personality (Back et al., 2010).

Without framing personality in terms of the Big Five model, but centering attention on features of people's organizational personalities and emphasizing the sociologist perspective[37], Burt et al. (1998) show that personality varies in the presence of structural holes. Similarly, personality was also shown to vary with network closure (Kalish & Robins, 2006). Furthermore, applying the mechanism named PCO – Propensity to Connect with Others, Totterdell et al. (2008) found that PCO was strongly related to network size. The authors measured the relationship between social network characteristics and personality[38], given people's propensity to connect with others – making strong ties, weak ties, and joining others (bridging ties).

In another vein of investigation, scholars have been developing relevant work on the understanding the dynamics of human emotions in social networks (e.g., Totterdell et al., 2004; Fowler & Christakis, 2008; Tang et al., 2011). This line of research aims to understand how emotions penetrate people's social networks. These studies undertaken by researchers from different areas are based on the notion that social networks have

---

[37] Burt (1998, p. 64) says that "Personality as a concept seems to be no more popular with psychologists than sociologists, but the exchange between sociology and psychology in organization behavior focuses attention on individual differences above and beyond differences attributable to network structure.".

[38] In spite of the fact that PCO is not framed in terms of personality it is similar to the measurement of extraversion.

their building blocks in dyadic relationships, but also in the matrix of relationships surrounding each person.

In this sense, emotions (e.g., surprise) and more enduring states (e.g., happiness) have been studied in both real and virtual worlds (i.e., Internet). Both emotions and happiness (or sadness) are feelings which have been proven to be transmitted in social networks. And both of them concern *affect*, which refers to a range of feeling states, including different moods and emotions (Totterdell et al., 2004).

Neuman & Strack (2000), show that when individuals are provided with different plausible causes for an affective response of unknown origin different emotions are activated. The findings have indicated that affective feelings can be transferred between people through a mechanism of mood contagion, and the other's emotional expression is sufficient to automatically evoke a congruent mood state in the listener. It should be noted that to achieve this effect it is not necessary to use verbal or semantic information about the emotion of the target person or an emotion-elicitation.

Totterdell et al. (2004), report that employees' feelings depend on the network of people with whom they work. More specifically, the feeling of affect shared within a group of employees is a predictor of affect towards other employees in the network, if the similarity of their structural position is taken into account. Equally relevant is the finding that the affect determined the network structure, rather than the other way around. The authors advocate that affect might have determined who people chose to work with.

Fowler and Christakis (2008), in turn, have found that happiness it is not only an individual experience or an individual choice, but that it is a property of groups of people. This seems to agree with the so called "affective revolution" of Barsade et al. (2003)[39]. Thus, happiness can be transmitted to others in a network flow that can reach three degrees of separation. The happiness of someone is associated with the happiness of others, who are located nearby or in other clusters of happy people. Furthermore, this effect holds true in both the real and virtual worlds (Whitfield, 2008).

Despite this work, a certain downplay of the role of psychological attributes in the social network analysis has been commonly accepted. Taking a macro level

---

[39] In the affective revolution, feelings are not understood as a solely personal experience, but as result of how people socially share and influence each other's affect at work and how this affects work life.

perspective of behavior, the assumption is that all humans act in the same fashion. Human behavior is seen here like nodes of a network that are represented by a homogenous set manipulated by environmental influences alone (Crosier et al., 2012).

In another line of research, scholars have been concerned with the role of a specific emotion (i.e., surprise) and its relationship with the information contained in a piece of data, or with the cognitive perception of novelty.

The concept of information in the context of its quantification has been largely debated in literature (Aczel & Daroczy, 1975; Cover & Thomas, 1991), particularly since the work of Shannon (1948). The debate ranges from the quantification of the information included in a piece of data, to the measurement of the information yielded by one event (Cover & Thomas, 1991). Another perspective on the concept of information is the fundamental effect that a piece of data has on an observer by replacing their prior beliefs with posterior beliefs. Deviation measurements between prior and posterior beliefs can be considered an aspect of surprise information (Baldi & Itti, 2010). In regard to novelty, scholars report that surprise is a specific consequence of the appraisal of novelty (Finkenauer et al., 1998). It measures the improbability or novelty of a certain event (Strange et al., 2005), as detailed in the second chapter.

In summary, similarly to the research into weak ties, many empirical works have debated Burt's conclusions alluding to the benefits and constraints related to the existence of structural holes. Usually, the viewpoint explored by the majority of those works relies on the bridging factors identified by the weak connection to socially distant groups or by the structural position of the sender of information. Common to all these theories is that they report on the delivery of novelty based on its assumptions about bridging factors. However, literature has not explored whether the delivery corresponds to the perception of novelty by the receiver – the other side of the bridge – neither whether centrality roles couple with the receivers' choices when they select information. This means, how their selection corresponds to the delivery of information when this is determined by central nodes. On the other hand, as far as I know, there is no prolific research on how to join social interaction and emotional reaction in order to apply it to digital media systems.

Motivated by these questions and with the aim of presenting a new approach to accessing social network data to avoid the effect of social echo chamber in personalization of Web-based applications, three empirical studies are presented below. These studies are based on the same survey and sample of participants. The next chapter discusses the first one.

CHAPTER 6

SURPRISE AS A PROXY OF NOVELTY

6.1 Overview

Data about strong ties and similarities between individuals from social networks have become an important resource to personalize Web-based services. Some authors have previously pointed to constraints related to Web personalization due to the diminishing diversity of viewpoints within communities. This is related to the lack of novelty of information shared. Structural bridges may be an accurate source of social data to introduce novelty on the receiver's side. It may create a new kind of data source for personalization. However, literature often debates the delivery of novelty but not its perception. This study proposes an alternative method that uses surprise as a proxy of the perception of novelty. It introduces a new approach to investigate the bridging process and how to confirm bridging assumptions. The results point out solutions for some constraints identified in digital media systems. A sample of 56 individual emotional responses to content selections in a social network environment is analyzed. Multivariate regression analysis shows that both weak ties and non-redundancy are predictors of surprise, but not all non-redundant structural holes identified are related to surprise. This attracts attention to the generally accepted bridging assumptions. It contrasts the differences between novelty delivered and perceived. Furthermore, socially distant ties and emotional support (closeness) play a relevant role in this regard, as well as the number of strong ties in the triads that surround structural holes. This method can potentially be useful in empirical work where novelty or its underlying dimensions are used (e.g., novelty vs. creativity).

6.2 Introduction

In order to engage their public in more effective and striking ways, digital media entrepreneurs are using data from social networks to personalize Web-based services (e.g. searching and recommendation).

Those approaches have been based mainly in the strong ties and homophily processes. Instead we have chosen to test a new approach through data derived from network bridges. The basic principle behind this approach is the idea that a network bridge can be used to establish the delivery of novelty (Borgatti & Lopez-Kidwel, 2011), avoiding the use of redundant information for the end-user.

Two different theories (Granovetter, 1973; Burt, 1992) debate bridging factors in the scope of the delivery of novelty, but neglect to say whether this novelty is perceived by the receivers. This is particularly important because is the perception of novelty that could confirm if the bridges are effectively associated to novelty delivering.

To overcome this problem it is necessary to extend the current methodologies to deal simultaneously with content and users' cognitive reaction and secure that the content delivered by an identified bridge corresponds to the perception of novelty.

One major way to engage users with content is through the emotions raised when the information is perceived. To test this method, I have raised two hypotheses: a) surprise is an accurate proxy of novelty, and, b) bridging factors are predictors of a surprise response. Thus, when surprise is elicited and the bridging assumptions are verified, then, theoretically, we can assume that novelty is perceived.

Let me briefly develop on the concepts behind each of those hypotheses:

a) The adequacy of the proxy proposed is based on neuroscience and psychology studies, which confirm that surprise accompanies novelty (e.g., Berlyne, 1960), and despite the fact that surprise can be elicited in events not related to novelty (e.g., Barto et al., 2013), surprise is the triggered emotion when novelty is perceived (e.g., Strange et al., 2005)[40];

b) This study through the assessment of surprise response, tests whether both assumptions on bridging factors predict surprise, and if each one of the assumptions as a factor to deliver novelty corresponds to the novelty perceived.

Moreover, as modern sociological theory suggests, novelty is found through weak ties that span structural holes. This raises the question of whether both bridging factors are coincident or correlated, when delivery corresponds to perception of novelty. To elaborate on these questions, I hypothesize that surprise is elicited either when the information is delivered by one single bridging factor or by the composition of both.

[40] This topic is not debated here more extensively because the extent in which surprise and novelty are interrelated it is largely debated in the second chapter.

On the other hand, considering the bridging measures introduced by Burt (1992), it is expected that for every bridge defined there will be a corresponding perception of the novelty delivered. Because scholars have not tested the veracity of the correspondence between novelty delivered and perceived, it is not possible to confirm whether this metric is accurate. The proposed methodology allows test this correspondence and debating this issue.

The aim of this study is twofold. Firstly, test a methodology that looks for the most efficient bridging factors in the context of novelty perception. That means analyzing Granovetter's (1973) theory, on the relevance of the socially distant ties and the most relevant dimensions to the tie strength construct, as well as considering Burt’s (1992) theory, on the redundancy of the connections spanning structural holes and observing the strength of the ties in the triads formed by sender, receiver and common connections with a third party.

Secondly, overcome the constraints associated with the effect of social echo chamber by showing the opportunities of applying data organized by bridging factors.

6.3 Bridging measures

As referred before I will examine which bridging factors meet the perception of novelty by the receiver. Assuming that novelty is confirmed by the emotion of surprise as its proxy, the strategy became to observe the emotional response of the receiver when accessing the contents.

Two kinds of bridging factors were tested. According to Granovetter (1973) a bridge appears between two individuals weakly tied, or, if strongly tied, they cannot have third-parties common to both.

In Burt’s assumption, a bridge exists when there is non-redundancy between the individuals connected through a structural hole. This implies the nonexistence of common third-parties between these individuals. To evaluate the structural holes and to define non-redundancy I have used triad-level measurements, instead of summary measurements (Kalish & Robins, 2006).

The following hypothesis will be explored:

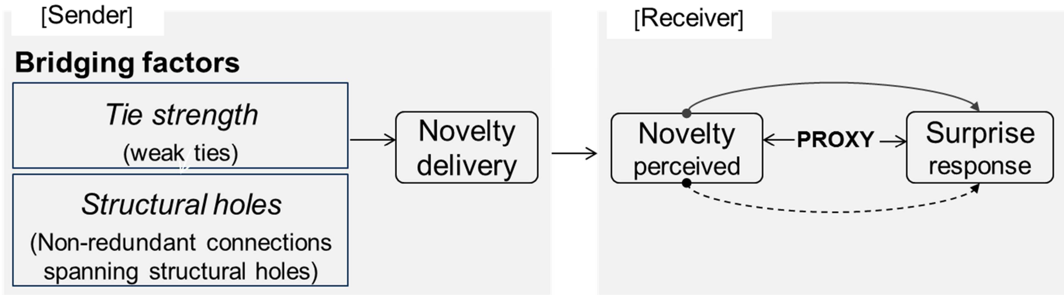
H1: *surprise is a proxy of novelty*;

H2: *surprise is elicited either when the information is delivered by one single bridging factor or by the composition of both.*

In order to test them, data was collected from a group of volunteer participants in a survey which was split into two phases.

The process of novelty delivering relying on the action of the sender of information is the main assumption behind the bridging mechanisms on. Though, it is important to note that a bridge determines both who delivers and who receives resources (e.g., new information). In this study, the participant whose content was selected is called sender and the participant who has selected the content, the receiver, is referred to as the selector of information. Participants shared and selected contents of other participants by privately describing the emotion they perceived, which includes surprise. "Surprise response" is the output of the action undertaken between the "sender" and "receiver", as Figure 1 shows. This work presents an approach to studying how structural conditions may explain the emotional reaction of surprise and provides an alternative method to control the whole bridging process.

Figure 1 – Conceptual model on surprise as a proxy of novelty.



6.3.1 Procedures

The two phases of data collection were undertaken in an online setting and by means of an online questionnaire. A Project's Facebook Page (PFP) was created as a platform for the participants interactions. In the *first phase*, participants shared content on the Project's Facebook Page (PFP) and have forwarded the selected posts to the message box of the PFP[41]. Participants were asked to register the emotion perceived whenever they accessed and selected a post.

Several procedures were created in order to avoid possible bias based on expectations or learning from others:

- Participants were not directly questioned about their perception of surprise to avoid biasing them by any kind of expectation or misconception concerning the real emotion perceived (Ramiller & Wagner, 2009);
- Participants classified the emotion perceived without knowing how others classified identical posts (they didn't have access to the classifications by others).

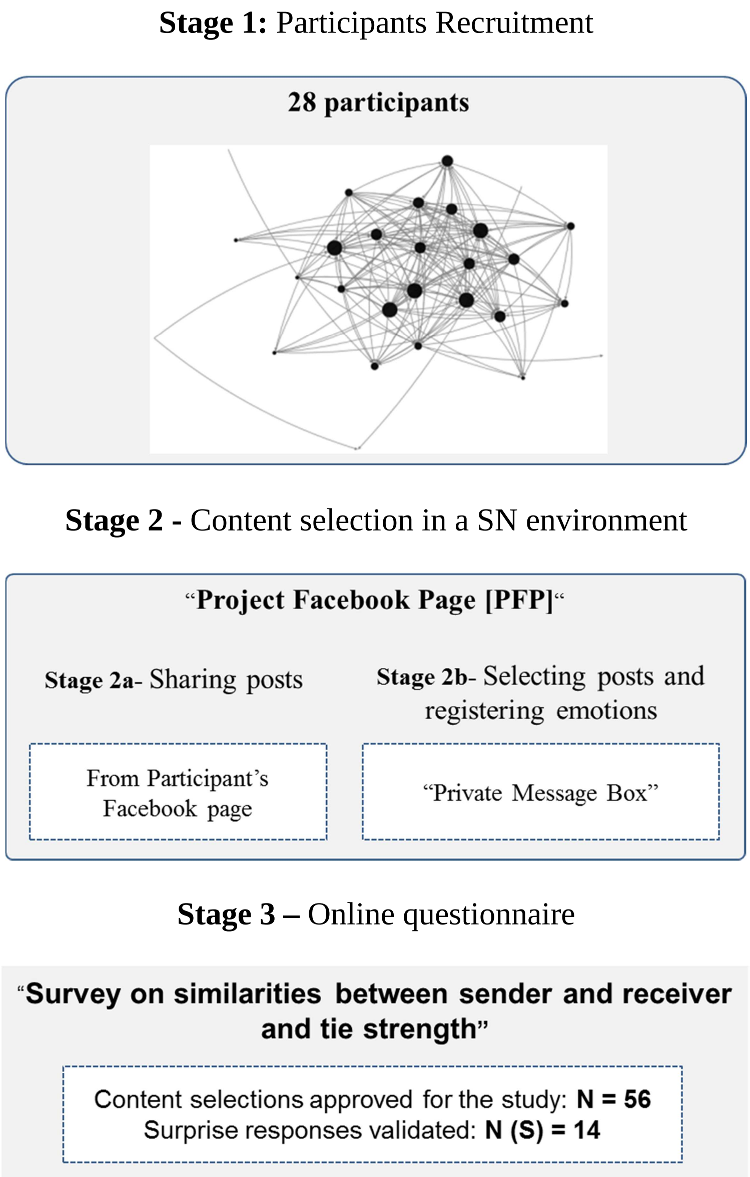
A list of emotions to classify these messages was previously distributed to the participants (Table 1). This first phase lasted five days. This phase includes stages 1 and 2 of the flowchart presented below (Figure 2). In the stage 2b), the number of content selections (posts selected) by each participant varied between two and four. As Figure 2 shows, in total, 97 content selections were validated, but this number was reduced to 56, because only the first and second selection of each participant was counted. The aim was to equalize the number of times that participants appear in the data and reduce the possible data bias.

In the *second phase* participants had to fill in an online questionnaire. The survey (Appendix A) is structured along two main topics, level of relationship and friendship perception. Firstly, they were asked about their perception of friendship with those participants from whom they picked posts to classify emotions. Secondly, they

[41] This description requires some knowledge about relevant Facebook functionalities, which we briefly review here. Nevertheless, it is important to highlight that the experiment undertaken in this dissertation, were not influenced by EdgeRank - the algorithm used by Facebook to determine what articles should be displayed in a user's News Feed.
Facebook main purpose from a user perspective is to become virtual friends with other users, and to communicate and stay informed about their activities and interests. When a user sends a friend request to another user and the latter accepts the request is established the friendships. It generates the so-called Facebook friend. Friends can usually read each other's contents ("posts"). Posts are unaddressed text messages, possibly enriched by photos or videos, which can be commented on and "liked" (by clicking a "like"- button). {In the survey this actions were not required, being participants even discouraged to do it}. Such posts appear on the users' "news feeds", a collection of friends' posts and notifications of other activities of friends (e.g. when someone changed his/her profile picture). Users can post on their own "walls" or on their friends'. {The timeline of the PFP was the main page used by participants in this study. Here, they shared contents and had access to other participants' posts}. Walls show all posts and notifications related to a certain user (whereas news feeds show posts of all of a user's friends). Users can also tag friends in their posts. This way, the post does not only appear on the user's wall, but also on the tagged person's wall. Friends' privacy settings and filter options set by the user determine which posts and notifications appear on news feeds and walls (Bohn et al., 2014).

were asked to indicate which participants they considered to be close friends, regardless of whether they picked their posts or not. A flowchart describing this process is presented below.

Figure 2 – Flowchart of the several stages of the survey.



Legend:

- Stage 1: Participants became members of the "project Facebook page". Two kinds of ties were designed by this survey. Friendship ties, as above shown and the ties established when the receptor selected the sender's post.
- Stage 2a: On News Feed: Published posts: 199; Selected posts: 174. Here, participants published posts from their own Facebook page and by following their

own interests. Each participant was invited to: a) view all the posts published; b) select the most interesting posts.

- Stage 2b: Links created: sender and receiver: 202. Participants selected the posts and sent it to the private message box of the PFP. They also registered the emotion perceived. The emotions registered were selected from a list of 10 emotions provided previously to the participants. The name of the participant who posted the post is also registered.
- Stage 3: The questionnaire answers were based on the names of the post authors selected – senders of the posts selected by the receptors – see stage 2b.

To confirm the option taken and avoid instability on the regression models, the co-linearity among content selections was analyzed to test their independence and so too was the non-co-linearity of the data. The agreement between content selections was analyzed with Intraclass Correlation Coefficient (ICC) using the two-way random model (McGraw & Wong, 1996; Shrout & Fleiss, 1979) and Cohen Kappa for nominal variables. Variables whose upper bound for ICC computation was above 0.50 were not considered, as suggested in similar literature to this field of work (e.g., Duncan & Raudenbush, 1998). The data for tie strength (ICC = 0.226) and surprise (Cohen Kappa = 0.103) show independence in its observations.

6.3.2 Sampling characteristics

Sampling procedure involved different processes of recruiting (direct appeal and using the 'snowball' technique[42]) in order to find people that know each other (living in the same university dorm) as well as people from other contexts. The aim was to ensure that the sample would not be formed only by random connections, or by connections only centered in the same kind of *foci* (participants from a dorm). I also intended to test different kinds of relationships (tie strength) and similarities between individuals. Hence, the sample should hold a reduced level of randomness, but still be representative of a large population.

[42] The most common methodology used in whole ('sociocentric') networks is the snowball sampling, commonly applied in small-to medium-sized networks (Wasserman & Faust, 1994).

56 emotional responses to content selections were validated in the study from 28 participants (16 males, Mean (M) = 19.7 years, Standard Deviation (S.D.) = 1.4 years, 12 females, M = 21.7 years, S.D. = 5.1 years).

Participants' age averaged around 20.5 years (S.D. = 3.52). Ninety-six per cent of them were between 18 and 23 years-old, the youngest was 18 and the oldest 37 years-old. The majority of the participants were Christians (79%, n = 22). The others were Buddhists, Muslims or Agnostics (21%, n = 6).

All tasks pertaining to the two study phases were performed online.

The sample set for this study consists in data from one group formed by seventeen dorm residents[43] and another by eleven non-residents (friends of friends). The reason for choosing those two groups was to force interactions within different kinds of connections as well as with the surrounding environment. Each participant was encouraged to invite up to five friends, and especially if those friends did not belong to the dorm or to that same university community. On the other hand, by recruiting in a dorm I expected to capture different levels of interaction, different kinds of relationships (tie strength) and similarities/dissimilarities between individuals. Moreover, with the individuals who were external to the dorm, playing the role of friends of friends, we aimed to extend the grades of separation from each recommender considered in the study and so to diversify the network.

6.3.2.1 Surprise elicited as dependent variable

Surprise is the dependent variable in the study. The registered emotions were coded as a dichotomous variable: surprise (n = 14) and not surprise (n = 42). The assessment of the emotions was done by using a scale previously delivered to participants[44]. Thus, each of the contents posted was rated by the selectors of information using the scale (see Table 1) for the emotion felt when the content was accessed. Participants were able to describe more than one emotion, either by mixing different categories of emotions or by mixing subcategories with categories. Thus, the

---

[43] At University of Texas in Austin (UT).

[44] In this study the "Differential Emotions Theory" (DES scale) (Izard, 1977, 1991) that postulates ten primary emotions was adopted and crossed with the sub-categories defined by Derbaix & Vanhamme (2003). Both include surprise, as shown in Table 1.

emotion of surprise was followed by another emotion in several situations, which was positive (e.g., surprise + joy) or negative (e.g., surprise + anger) (Ekman & Friesen, 1975; Meyer et al., 1994). Moreover, when the emotion perceived was "surprise", participants were asked to write down why he or she was surprised, as well as to describe other emotions that could complete their sense of surprise.

Table 1 – Emotional scale.

Emotions	Sub-categories
Surprise	Surprised, amazed, astonished
Enjoyment	Joyful, delighted, happy
Interest	---
Distress	Sad, downhearted, discouraged
Anger	Angry, mad, enraged
Fear	Afraid, scared, fearful
Disgust	Disgusted, feeling of distaste, feeling of revulsion
Contempt	Disdainful, contemptuous, scornful
Shame	---
Guilt	---

6.3.2.2 Tie strength and Redundancy as independent variable

The two independent variables that play the role of network factors are the tie strength and redundancy.

1) *Tie strength*: Several distinct types of social interaction were identified – for example, some participants spent time with other participants on a daily or weekly basis, and some do not feel comfortable to borrow money from others (see Appendix A).

Several methods have been used to construct the overall measure of social interaction since each person is potentially connected to another by several types of relationship[45], the most common variables quoted on literature were followed (for a revision see: Petróczi et al. 2007).

[45] Tie strength was measured according to the following weighing between variables (see indexes of the variables on Table 2): Tie strength = [(V1 + $V2_{\text{type of relationship}}$ + ($V2_{\text{private correspondence with}}$ * 2) + V3 + (V4 *

Granovetter's (1973) tie strength definition[46] was used. The argument of Marsden et al. (1984) on closeness[47] was also considered, as being the best indicator of tie strength. Equally, the emotional support to characterize strong ties relationships was identified, as reported by Wellman and Wortley (1990). Finally, the choices taken agree with Petrosky's (2011), which mention two dimensions to conceptualize tie strength: intensity and valence. The former is about frequency of contact and the latter refers to "the affective, supportive and cooperative character of the tie" (ibid. p. 44).

In this sense, I characterize tie strength through four variables, each of which is described by several survey items. Such variables are: intensity/ communication and reciprocity, intimacy, duration/ amount of time, emotional support, as detailed in the Table 2 below.

Table 2 – Tie strength construct.

Variable	Detail
V1 - frequency of contact and reciprocity	Tie strength scores were weighted both to distinguish the relevance of different items, as to differentiate variables between themselves[48]. E.g., *Frequency of contact and Reciprocity* was measured with two questions. The first was: "How often have you had contact with each person that you mentioned above". Responses were rated on a ten point scale, where 1 represented "other", 2 "once a year" and 10 "every day". The scale is not linear to emphasize relevant differences. Thus, the second less quoted answer "twice a week" was rated with 7. Similarly, the same procedure was used with the other variables. The second question assigned to this variable was a request to write down the names (four of them) of other participants that he/she knew best (having met them socially/ professionally, e.g., in sports, parties, work, classes). Participants were rated with a score of 5 if the answer was reciprocated by the other participant, and with a score of 3 if not.

2)] / 7. The scores of the weak ties could range from 1 to 4 and strong ties from 5 to 9. The final value of tie strength could range between 2 and 11, depending on the proportional strength set for each variable.

[46] As presented in the fifth chapter, Granovetter (1973, p. 1361) states that "the strength of a tie is a combination (probably linear) of the amount of time, the emotional intensity, the intimacy (mutual confiding) and the reciprocal services which characterize the tie".

[47] Closeness has been considered the best indicator of tie strength, such as underlined by Gilbert & Karahalios (2009) that assessed the strength of study participants' friendships in Facebook.

[48] Some variables and items are scored with a double weight. The variable "intimacy" is made up of two items. One of the items "having private correspondence with" was scored with a double weight. The "emotional support" items were also scored with a double weight. Moreover, some of the items were scored with a nonlinear scale. This was determined by the relevancy of the item for the tie strength concept. E.g., "How often did you have contact with each person?" If the answer was "every day" then 10 points were given and 7 or 7-n (n<7), for scores below.

Variable	Detail
V2 – Intimacy (confidence)	The *intimacy* was measured through two more questions. The first one on “What type of relationship do you have with the people that you mentioned above”. Participants were rated on a five-point scale, ranging from 0 ("Acquaintance / Other - Not close at all") to 5 ("Partner, Boyfriend/ Girlfriend"). The second question was “Who are the people that you mentioned above with whom you have private correspondence”. Participants were free to select up to four names. The selected names were marked with “yes” and got a 5 points rating, if not they got a 0 points rating.
V3 - Duration/ amount of time	The *duration/amount of time* contained a single question: “Indicate for how long you have known each of the mentioned people” and was rated on a three point scale, where 0 represented "other", 2 “less than three months” and 4 "more than one year".
V4 - Emotional support	The *emotional support* is a construct of three questions. The first one, was with whom the participant felt familiar enough to ask “to borrow a small sum of money from”, and the second one was who the participant would contact if “feeling sick, or needing health support”. Both were rated on a three point scale, where 0 represented "no", 2 “uncertain” and 5 "yes". The third and last question was about “how close do you feel with” the four participants from whom he/she picked content, and was rated in a five-point scale, ranging from 0 ("Don’t feel close at all"), 2 “I don’t feel very close” to 5 ("I feel very close").

2) *Redundancy*: To measure bridging factors applying Burt’s (1992) theory, I needed to evaluate the degree of redundancy between the participants that span structural holes. Thus, triadic-level measures were applied (Kalish & Robins, 2006). A triad is a set of three persons that tend to close through a third person, forming a triadic closure, due to propinquity or cognitive processes (Goodreau et al., 2009), in which the strength of the ties among individuals plays a determinant role. Propinquity represents the process in which two people encounter due to the time shared with a third. Cognitive processes, highlighted by the social balance theory (Heider, 1958), are represented by cognitive events in which two people may appreciate each other mutually by their agreement on a third person. Hence, even if two individuals share distant ties, they may share similar perspectives and access similar information. This fact may preclude the novelty between the two individual ties when they share information.

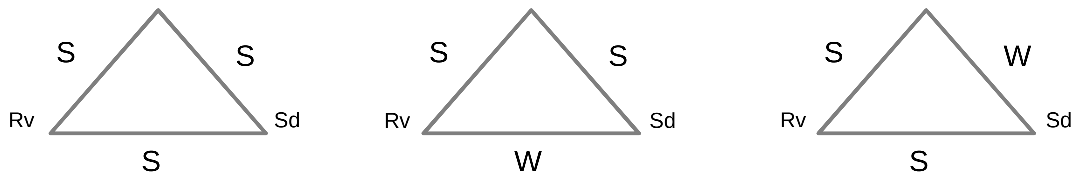
To evaluate the existence of triads in the data, first, the existence of common connections between each pair sender – receiver was examined. Then, the ties among the individuals included in each triad formed were analyzed: tie strength or absence of ties, as explained in the “forbidden triad” role of Granovetter (1973). To accomplish that, the following information from the questionnaire was used: (1) tie strength, as described above; (2) information from the question “names of the participants that you know best (people that you have been meeting socially/ professionally, e.g., in sports, parties, work, classes)”. In order to gather the maximum information possible on ties among participants, the collected data in the survey was confronted with the data on “friendship” ties downloaded from the project’s Facebook page through the NodeXL software (v. 1.0.1.210).[49]

Table 3 – Types of triads and strong ties per triad.

	WSW/ WWS	WSS	WWW	SWW	SSS	SSW/ SWS	<=1 ST	>1 ST[50]
Nr. of pairs sender-receiver per kind of triad	13	7	9	12	10	18	45	27
Mean (nr. of triads/connection with triads[51])							1.70	2.19
SD (nr. of triads/connection with triads)							0.95	1.24

Six types of triads were considered (Table 3). Two of them represent network closure: SSS and SSW/SWS (representing strong network closure).

Figure 3 – Closed Triads



“Rv” means receiver and “Sd” sender. “S” stands for strong tie and “W” weak tie. The first letter means the tie between sender and receiver (the selector), the second letter the tie between sender and third-party, and the third one the tie between receiver

[49] We tested all the results for redundancy, with and without those forbidden triads forming four-cycles. The existence of four-cycle indicates that structural holes are not present. The forth element may induce a closure of the cycle through the third-party. I tested all results for 4-cycles redundancy and I found out that the increment of redundancy, through a common fourth element, did not change the statistical results found for the variable redundancy. Thus, I do not include 4-cycles redundancy results in the discussion.

[50] ST – Strong Tie.

[51] 27 connections sender-receiver with <=1 ST, 27 with >1 ST and 29 without any triad.

and third-party (e.g., WSW: Weak-Strong-Weak). The closure of a triad WWW (representing a weak network closure) was not considered because such connections include distant ties (up to 50% in average in the study, as shown below), thus, there is no redundancy. From the six types of triads, four represent the existence of structural holes: WSW/WWS, WSS, WWW and SWW.

6.3.3 Results

The introduction of receptor's perspective to confirm the validity of a bridging probably changes the number of bridges counted by comparison to the result obtained by using the usual bridging factors approach. This section helps to confirm this and observe how structural network conditions may predict the emotional response. The identification of bridges and the corresponding emotional responses, in particular surprise, provides a larger control of the whole bridging process. It informs about receivers' individual characteristics and creates new valences of observation, which can be used, for example, to compare them with those of the senders. This may allow moving forward in the examination of predictive factors in the delivery of novelty using data from social networks.

6.3.3.1 Tie Strength

The Table below presents the separate scores obtained for the ties coded as weak and strong and for the final value of tie strength between sender and receiver.

Table 4 – Descriptive statistics based on the scores of the variables that characterize tie strength.

	V (1; 2; 3; 4)	N	Minimum	Maximum	Mean	Std. Deviation
V1 - Intensity/ communication and reciprocity	weak tie	33	2	4	4.121	1.727
	strong tie	23	5	16	8.652	2.707
V2 - Intimacy	weak tie	33	2	8	2.727	1.484
	strong tie	23	3	10	8.391	2.407
V3 - Duration/ amount of time	weak tie	33	1	4	1.969	1.103
	strong tie	23	3	4	3.565	0.506
V4 - Emotional support	weak tie	33	2	4	4.181	1.590
	strong tie	23	5	9	11.043	2.915
Tie Strength	weak tie	33	2	4	2.636	0.822
	strong tie	23	5	9	6.782	1.412
	tie strength	56	2	11	5.607	3.061

Descriptive statistics for variables used in the analyses of tie strength are presented in Table 4, which describes the contribution of each variable (V1, V2, V3 and V4) for the tie strength's construct. It also shows how each variable contributes to the score of weak and strong ties.

V4 – Emotional Support is the variable with the highest rate to characterize the strength of the tie (Mean = 7.000; SD = 4.058), but also the one which received lower values when the participants did not know each other. This variable emphasizes the differences between ties (weak and strong) and the score range of the weak ties that vary from 1 to 4 and the strong from 5 to 9. The final value of tie strength could range between 2 and 11, depending on the proportional strength set for each variable. This draws the attention to the existence of "socially distant ties" between the weakly tied participants.

A socially distant tie means a tie between two individuals (sender and receiver) that had never had any contact before the study or almost nonexistent contact. The score classifying weak ties varies from 2 to 4, with 57.5% scoring 2. Surprise response was elicited 42.9% for a tie strength with a score of 2 (n = 6); 29% when scored 3 (n = 4) and 14% when scored 4 (n = 2). Finally, surprise response was elicited 14% (n =2) for a strong tie.

To summarize, given the results above, "socially distant ties" are the weak ties scored 2. These ties were also rated with low values in the variable of Emotional Support (V4). Weak ties which scored 3 or 4 are present in equal percentages (21.2%). Given the relevance of the Emotional Support variable to determine tie strength [52], and the findings related to its scores (V4), this variable seems to be an accurate dimension to detect socially distant ties. Moreover, almost half of the surprise responses (42.9%) are related to the lower scores on tie strength. Thus, it seems that surprise is mostly elicited by weak ties from people who are socially distant.

6.3.3.2 Redundancy

The Table below shows the total number of triads when there is fewer than or equal to one strong tie (<= 1 Strong Tie) and more than one strong tie (> 1 Strong Tie). The results are split as a function of the redundancy state and the average number of triads between each connection.

Table 5 – Number of triads between sender and receiver.

Ties between sender a selector			<=1 Strong tie in the triads				>1 Strong tie in the triads			
	Redun-dancy	N (Tie & Redund.)	Triads (T<=1ST)	Mean	SD	Max.	Triads (T>1ST)	Mean	SD	Max
Weak tie (N_{WT} = 33)	0	26	23	2.18	1.25	5	--	--	--	--
	1	7	8	1.75	0.50	2	7	1	0	0
Strong tie (N_{ST} = 23)	0	3	1	1.00	--	1	--	--	--	--
	1	20	14	1.27	0.46	2	52	2.60	1.88	5
Weak tie & Surprise (n = 12)	0	11	5	1.50	1.00	3	--	--	--	--
	1	1	--	--	--	--	1	2.00	1.41	3
Strong tie & Surprise (n = 2)	0	0	--	--	--	--	--	--	--	--
	1	2	--	--	--	--	2	2.00	1.41	3

The results show that among all bridges that match the assumptions of Burt (1992) related to non-redundancy (N = 29 [N_{WT} = 26 + N_{ST} = 3]), only eleven are related to novelty perception ($N_{WT\ \&\ Surprise}$ = 11). On the other hand, the prevalence of the number of strong ties on triad formation determines the existence of redundancy.

[52] Which agrees with Marsden et al.'s (1984) argument on "closeness".

Table 5 shows that the non-redundancy is verified (Redundancy = 0) only when the triad is formed by a maximum of one strong tie (<=1 Strong tie). This is verified regardless of whether the tie strength between sender and receiver is weak ($N_{Tie\ \&\ Redund.}$ = 26; $T_{<=1ST}$ = 23) or strong ($N_{Tie\ \&\ Redund.}$ = 3; $T_{<=1ST}$ = 1). Whether the triad has more than one strong tie, redundancy is verified, which confirms that redundancy is determined by the number of strong ties in the triad.

On the other hand, surprise responses are mostly related to non-redundancy and weak connections ($N_{Tie\ \&\ Redund.}$ = 11). In fact, surprise response related to redundancy (Redundancy = 1) is only verified in three cases: one, in a weakly connection ($N_{Tie\ \&\ Redund.}$ = 1; $T_{>1ST}$ = 1) and two, in a strong connection ($N_{Tie\ \&\ Redund.}$ = 2; $T_{>1ST}$ = 2), in which the triads have more than one strong tie (>1 Strong tie).

Table 6 – Pearson correlations between triads and bridging factors.

Variables	(1)	(2)	(3)
Triads			
(1) Triads <=1 Strong tie	--		
(2) Triads >1 Strong tie		--	
Bridging factors			
(3) Redundancy	0.037	* X2 (1) = 56.00	--
(4) Tie strength		* X2 (1) = 25.77	*X2 (1) = 25.77

* p < 0.001

Findings confirm strong evidence of a relationship between triads with more than one strong tie and redundancy (X2 = 56.00, df = 1, $p < 0.001$) as shown in the Table 6, where all closed triads of the study are related to redundancy. There is also strong evidence of a relationship between these triads and tie strength (X2 = 25.77, df = 1, $p < 0.001$), where 87% of the strong ties are related to redundancy, while only 18.2% of the weak ties are present in such triads. This shows that triads related to redundancy are predominantly dominated by strong ties.

Finally, it is worth examining the answers issued from participants to describe, in their own words, why they were surprised when this emotion was picked out by them. Nine of them answered “expectedness”, one referred “novelty”, two referred “new perspectives” and two did not provide any answer. These answers are in line with what has been reported in published literature on surprise and novelty, as mentioned above in this work.

6.3.3.3 Bridging Factors and Surprise

In this section the associations between the bridging factors and surprise are examined. In order to test the two hypothesis raised, firstly, the relationship between surprise response and both bridging factors was analyzed. This was undertaken through logistic regression analyses predicting surprise using tie strength (Granovetter, 1973) and redundancy (Burt, 1992) separately as independent variables (Table 7).

Table 7 – Coefficients from regression model predicting surprise and redundancy.

Predictors	Redundancy	Surprise
Bridging factors		
Tie strength	5.477 (.001)	-.408 (.030)
Redundancy		-.125 (.012)

* Applying Granovetter’s (1973) forbidden triads.

To test Granovetter’s (1973) assumptions I analyze the association between surprise and tie strength. The results suggest that there is a significant positive relationship between surprise and weak ties (36% weak ties vs. 9% strong ties) and the odds of being surprised decreases when the tie is strong (odds = 0.408, p = 0.030, 95% Confidence Interval (CI): [0.182, 0.915]). Thus, weak ties are determinant to explain surprise, but it is important to note that distant ties (scoring 2 in the tie strength range) represent 42.9% of the weak ties for surprise responses.

Next Burt’s (1992) assumptions were tested and I have analyzed whether surprise is predicted by the independent variable of redundancy. The results tell that there is a relationship between these variables and that the odds of being surprised decreases with the redundancy (odds = 0.125, p = 0.12, 95% CI: [0.025, 0.630]). When considering the number of strong ties per triad in non-redundant connections (between sender and receiver), we observed that 62% of these triads have no strong ties, or just one strong tie (38%). Among these non-redundant connections, surprise is elicited in 61.1% of the triads that do not hold any strong tie and in 36.4% of the triads with one single strong tie. Considering redundancy, surprise is elicited in 11% of the triads with more than one strong tie. Therefore, the set of results obtained suggest that the prevalence of surprise response at structural level is strongly associated to the number and strength of ties forming the triads that surround the connection between individuals, separated by a structural hole.

Hence, the outcomes validate the hypothesis that surprise is a proxy of novelty. Thus, bridging factors are predictors of surprise response. This is verified for both assumptions on bridging. H1 is confirmed.

Secondly, by computing Pearson correlations (Table 6), evidence was found of the relationship between tie strength and redundancy ($X2 = 27.77$, $df = 1$, $p < 0.001$). The odds of experiencing redundancy increased whenever the sender and the receiver were strongly tied (odds ratio = 5.477, $p = 0.001$, 95% CI [2.585, 11.605]) (see table 7). This is evidenced by the fact that 90% of the weak ties are related to non-redundant connections, while only 23.1% of the strong ties are related to non-redundancy.

Next, the second hypothesis was tested. The key question now is to what extent is there any correlation between both bridging factors when there is a delivery of novelty. Both weak ties and non-redundancy were shown earlier to be predictors of surprise, so both are associated to perceived novelty. Concurrently, there is strong evidence of the relationship between them (weak ties and non-redundancy). It would seem then, that bridging factors could be correlated. Thus, this seems to justify the hypothesis that surprise is elicited either when the information is delivered by one single bridging factor, or by the composition of both.

In this sense, and given the strong associations between tie strength and redundancy, a multivariate regression was applied with forward stepwise selection of variables[53]. When computing surprise with each of the bridging factors, it shows an association with both strength and redundancy. However, when seen together, the redundancy remains statistically associated with the surprise response, but the tie

[53] We applied multivariate regression with forward stepwise, in order to estimate whether both independent variables, tie strength and redundancy, could predict surprise together.

strength does not. The odds of experiencing surprise decreased for higher redundancy (odds = 0.125, p = 0.012, 95% CI: [0.025, 0.630]). Therefore, the hypothesis confirms that surprise is elicited by each one of the bridging factors, but not by its correlation to predict surprise.

6.4 Discussion

Social network literature has mainly reported the existence of bridges delivering novelty, but only considering one single perspective: the information sender. It means that people who are socially distant, and located in previously separated groups (Granovetter, 1973), or connected by non-redundant structural holes (Burt, 1992) can receive novel information. However, these theories do not claim that the received information will be perceived as novelty. In this regard, I have shown that the receiver's perception of surprise plays a relevant role to explain and confirm the full process of bridging. The key question now is to evaluate to what extent this method can be regarded as a better approach to confirm delivery of novelty than those of other scholars, or even to confirm the theoretical assumptions underpinning the bridging mechanism of the two known theories.

It was shown earlier that surprise and novelty differ in their typical functions at neuronal level. While novelty is based on memory and on cognitive processes, surprise is based on expectations of systems capable of predicting. Furthermore, it is recognized that surprise accompanies novelty (e.g. Berlyne 1960) and psychology studies underline that surprise is the emotional state related to the evaluation of novelty (e.g. Smith & Ellsworth, 1987; Finkenauer et al., 1998; Strange et al., 2005). Some scholars state too that it is not accurate to say that surprise is always associated with novelty, but it is correct to claim that novelty perceived is always followed by a surprise response.

Following the above analysis, the method proposed helps to find an explanation for the events related to surprise, though not all events are related to bridging assumptions. In fact, not all surprise responses match the assumptions of network bridges.

In total, 14 surprise responses were reported by participants. Observing them using Granovetter's bridging factors, surprise is related to 12 bridges of weak ties and 2 of strong ties. It means that two receivers reported surprise, but they had a strong connection with the receiver (one scoring 5 and other 7 on tie strength, and 9 and 12 on

Emotional Support, respectively). In both cases participants had a feeling of unexpectedness, as can be seen by their justification for the emotion elicited[54] (see[55]).

When considering Burt's bridging factors, the novelty perceived is related to 11 bridges formed by non-redundant structural holes and 3 redundant connections. Thus, in some cases, the surprise response is elicited but they do not match the assumptions of the delivery of novelty defined by each author. Given this and the fact that novelty perceived is always accompanied by the emotion of surprise, as claimed by scholars, it seems adequate to consider these surprise responses as outliers in the context of bridging.

Considering that this method applies when the bridge meets the assumptions of the delivery of novelty, then the surprise elicited by the access to content corresponds to the novelty perceived. Moreover, both bridging factors – weak ties and non-redundant structural holes – are predictors of surprise. Thus, this method seems to be adequate to confirm the delivery of novelty based on its perception and to find the bridges that match the assumptions. On the other hand, these results are related to the fact that a specific emotional state can be predicted by specific structural conditions and determined by the rapport between pairs. This is in spite of the fact that in this study nothing has been said about the psychological attributes of each individual and how they may be related. This subject is debated in chapter eight.

Granovetter (1973) does not clarify how to distinguish between weak ties that act as bridges and others that do not, but this method may be helpful to specify which weak ties present better conditions to act as bridges considering the perception of novelty. As a matter of fact, regarding the weak ties, the socially distant ties are the ones that play the most relevant role in the delivery of novelty (42.9% of the surprise responses among weak ties are scored with 2 in the tie strength construct), and the emotional support (closeness) is the variable that best characterizes tie strength. Thus, two dimensions should be highlighted to distinguish the weak ties from the most

[54] Participant X: "Surprise, I was surprised because I didn't expect to feel this relaxed when listening to this"; Participant Y: "I felt surprised because the thumbnail looked like a grown up but it's actually a boy. It is a pleasant surprise because it's funny." In both cases the content accessed was an image. In fact, in most cases the contents associated to surprise responses in the study are images or videos. This might be related to the participants' age, which on average is 20.7 years old.
[55] See APPENDIX C – INSTRUCTIONS TO STUDY PARTICIPANTS - PHASE 1, ITEM 3.I.

accurate ones to match with bridging assumptions: a) social distance of the tie and, b) the emotional support between sender and receiver.

Burt (1992) has drawn out assumptions defining the best conditions for bridging actions. It is assumed that non-redundant structural holes can act as bridges. In this regard, the results of this study show that the number of strong ties included in the triads is determinant to identify structural holes related to surprise response, which excludes some bridges defined by Burt's assumptions. In fact, 11 out of 29 non-redundant structural holes are related to surprise, confirming that only in 11 content selections the receiver received novelty in a structural condition that avoids redundancy. It means that despite the number of structural holes between participants that exchanged content, and who do not hold a redundant connection with common friends, only 11 are related with surprise responses, which reduces considerably the number of bridges supposedly associated to the delivery of novelty. At the same time these findings show evidences of the relationship between redundancy and tie strength. Furthermore, the results suggest that non-redundancy is more prevalent in bridges delivering novelty (eliciting surprise) when the actors are weakly tied: 85.7% of the ties (n = 11) are non-redundant and weak.

To sum up, the findings show two important conditions for the perception of novelty that combine with the mechanism of novelty delivery. One concerns the number of strong ties in the triads to define redundancy. The other regards the distant ties, instead of weak ties in general, to define the tie strength with a higher probability of acting as an accurate bridge. And finally, a significant aspect to define distant ties should be associated with a low level of emotional support. This method, therefore, confirms which bridges correspond to the perception of novelty and are related to the stimulus of surprise.

Lastly, the findings show that non-redundancy is the bridging factor in the prediction of surprise that remains in the regression model when it aggregates all the variables under study. This means, that though both bridging factors are predictors of novelty they do not show this behavior when looked at simultaneously. The findings, at first, seem to agree with McEvily et al. (1999) who assert that there is no correlation between weak ties and non-redundancy. However, contrary to these authors, this study did not consider the infrequency of interaction as the single variable of coding of the tie strength, which may change the correlation between both variables. In fact, by eliminating the variable "Emotional Support" in the study, several strong ties became weak ties. Second, it should be noted that the correlation, or overlapping, between the bridging factors was not statistically proven. This may be due to a possible lack of sufficient statistical power, given the dimension of the study sample. This should be mentioned because the corresponding association between the variables seems to exist. In fact, non-redundancy is more prevalent when in association with weak ties, as mentioned above. Therefore, it seems accurate to say that the bridges confirmed by the method proposed – using surprise as proxy of novelty perceived – match the best conditions for delivery and reception of novelty, and that bridging is an important structural condition to explain the emotional reaction of surprise.

Last but not least, the method being tested in this study provides an alternative method to control the whole bridging mechanism, which could also be usefully applied in other studies on novelty or its underlying problems (e.g. novelty vs. creativity). When confirming the delivery of novel information in their studies, some social network scholars often do so by identifying other dimensions that are supposed to be a condition of novelty delivered (e.g., knowledge, innovation, creativity). Thus, they verify novelty as an underlying proxy to these dimensions. This is the case of Aral and Alstyne's (2011) work, which contends that strong ties are beneficial in network structures rich in structural holes. These ties create dense information flows that improve the access to novelty. These authors report different results from the ones analyzed here about the relevance of tie strength (weak ties) and structural holes. The reasons for those differences seem to be centered on the type of framework used. They rely strongly on studies that do not consider the reception of novelty, but on other dimensions thought to be related to the delivery of novelty, e.g., knowledge transfer (Hansen 1999; Reagan & McEvily, 2003), innovation (Staber, 2004; Obstfeld, 2005), and creativity enhancement (Fleming et al., 2007; Sosa, 2011). This method applied in these studies could extend its results, namely by allowing an accurate association between novelty and the dimensions mentioned above related to personal and/or cognitive performance.

6.5 Conclusion

There is not much published literature on empirical work regarding information flow through network relationships to validate network effects. This work tries to contradict this trend. In order to do this, I tried to clarify which bridging factors hold a stronger association with the perception of novelty by a receiver. The study deepens the understanding of the bridging mechanisms that are relevant for the delivery of novelty in a process of information sharing. This work also identified which individuals in a network, acting as senders of information, may play the role of brokers. A broker may intervene in the cognitive behavior of the receivers by suggesting new perspectives through the surprise effect. This seems to be a relevant contribution for the SNA field, as well as for the digital media field in regard to the problematic of the 'social bubble'.

This study presents a recognizable output, i.e., novelty perception, of the bridge mechanism and a new perspective over dyadic and network interactions surrounding these structural bridges. This is particularly useful to develop predictive models for these specific types of bridges. In fact, the prediction of novelty represents a potential solution for some digital media constraints, such as the 'echo Chamber' effect in the personalization of Web-based services (Sunstein, 2009), the 'Portfolio Effect' (Groh & Ehmig, 2007), identified in recommendation systems, and the effect of social echo chamber related to the current use of social data, as detailed in this dissertation. Such effects are related to the lack of diversity in users' viewpoints (Vargas & Castells, 2011), and, thus, a lack of novelty in information delivered (Golder & Yardi, 2010).

Regardless of the constraints and difficulties in keeping participants strongly engaged in long-term studies, it could be useful to extend this study to a larger population so as to reinforce or bring further clarifications on some analyses developed in this work. This research may also have faced some boundary constraints. Several ties were certainly out of the observation range, but I do not expect this uncaptured data would have interfered with the redundancy encoding results, to the point of observing a significant change in my conclusions. To support this claim I point out that an increment of non-redundant connections did not lead to notable changes in the 4-cycle redundancy tests described above.

Beyond the findings of the presented study, two questions remain unanswered: how is the information flow be influenced by the surprise effect? And, what type of relationship exists between the selection of contents and the friendship/ spatial proximity between senders and selectors of information (receivers)?

Finally, when aiming to extend the knowledge about the information sharing process in a social network environment, it is also important to analyze to what extent the centrality measures (Freeman 1979) interfere or may predict the factors behind surprise response and compare such results with the ones obtained with bridging factors.

The study presented in the following chapter looks for answers to these questions.

CHAPTER 7

STRUCTURAL HOLES AND SURPRISE IN CONTENT SELECTION IN SOCIAL NETWORKS

7.1 Overview

Limited attention has been paid to the influence that social network dimensions associated to senders position relative to the receiver may have on an individual's choices of contents. Thus, it is relevant to know how network dimensions (i.e., network centrality, structural holes, and tie strength) may influence the content selection by receivers. This raises the question of what determines such content selection. These relationships are empirically tested by using both social network data and participants' survey data. Findings show that despite the fact that degree and strength of tie are associated with central positions in the network, they are not related to individual's choices of contents. Findings also suggest that structural holes in association with the emotion of surprise, used as a proxy of the perception of novelty, offer a good representation of people‘s behavior when they select contents. These are valuable arguments to enhance content personalization with new perspectives for receivers.

7.2 Introduction

Social network literature is full of studies on bridging factors (i.e., weak ties and non-redundant structural holes), and centrality showing how each one determines the information flow (e.g., McEvily et al., 1999, Hansen, 1999; Holme & Ghoshal 2008; Kratzer & Lettl, 2008; Shi et al., 2013). However, despite differences in how they determine the information flow, they have only been studied by scholars from the sender's viewpoint. Having discussed in the previous chapter the perspective regarding network bridges, I am now going to analyze the benefits associated to network positions identified by centrality measures (Freeman, 1979). Particularly, it is examined the relationship between the network position, occupied by the sender of information, and the individuals' choices of contents.

The effects of central positions can easily be found in several online applications, such as on the delivery of a recommendation (e.g., the name of a book), which may follow criterions based on centrality (i.e., associated to the number of persons that bought or rated that book). The assumptions behind this approach are based on the knowledge that individuals look for staying in contact with popular items or connect with popular people.[56] Similar relevance seems to be attributed to the notion of social relevance.[57] Thus, the degree of centrality is often used to determine the value[58] of a network node (e.g., an online resource) located in such structural positions.

Recognizing that in many situations centrality is effective and provide good results, I would like to point to their weakest points and propose a more complex and powerful approach. The main problem of those methods is related to their over emphasis on a structural view and rationality neglecting the role of the elicited emotions when the receiver accesses the content.

On top of that, there is another reason why my approach can improve the traditional understanding that usually neglects information concerning the receivers. In fact, scholars have not been taking in account the receivers' role in the network. This means that is not clear how the structural position of the senders relatively to receivers may influence the individual’s choices of contents. In this regard, I analyze how network dimensions (i.e., network centrality, structural holes, and tie strength) may influence the choices of contents by receivers. I also evaluate if the content production of the sender, their exposure in the network, will influence such choices.

This raises the question of what determines the contents selection by receivers. What perspective should be followed in order to satisfy the target user? Should it be the sender's perspective or the receiver? Is there any difference between them?

Therefore, in this chapter I will examine the information flow between sender and receiver considering receiver’s content selection and their emotional reaction. I will explore how individual’s choices of contents are related with the perception of novelty.[59] The surprise response is applied as a proxy of the perception of novelty (Stiensmeier et al., 1995; Strange et al., 2005) to ensure that the bridging factors for the delivery of novelty correspond to the perception of novelty. In that way, I am certain

[56] Popularity in this context refers to people with a high value of degree centrality.

[57] From a network standpoint, social relevance derives from the high number of connections that a node has with adjacent nodes that also have a high number of connections. Thus, social relevance can be understood as the importance attributed to the assessment made by others about their choices concerning nodes connections.

[58] It assumes that individuals act opportunistically, calculating their potential benefits and costs.

[59] When this is verified, there is a non-redundant structural hole that connects sender and receiver.

that each choice of content corresponds to an emotion identified by the receiver (here dichotomized as ‘surprise’ and ‘no surprise’).

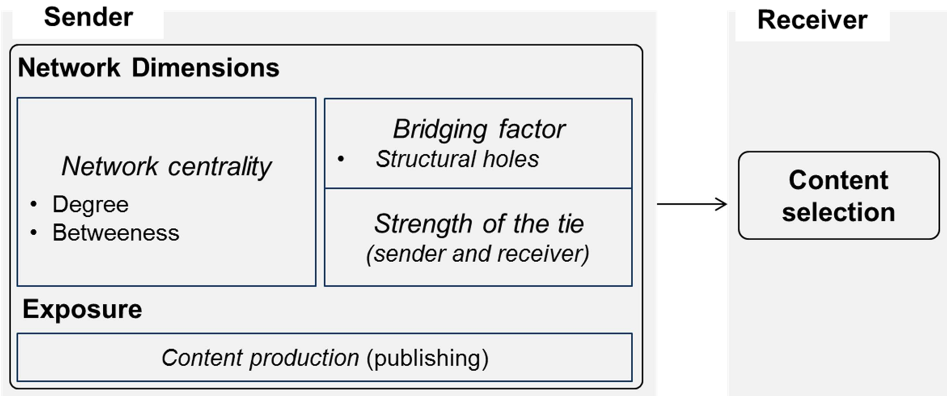
The analysis undertaken in the study presented will rely on two types of central nodes. One it is identified by its role in network bridges (Burt, 1992), and the other by two centrality measures, degree and betweenness (Freeman, 1979).

7.3 Central Nodes, receivers' content choices and hypothesis

The two types of central nodes considered in this chapter may have different implications to receivers. While centrality measures are based on the degree of its nodes (number of connections with adjacent nodes), network bridges are based on the ability to deliver novelty.

The approach proposed, presented in Figure 3, seems to be new in social network studies and introduces an important valence that contributes to complete the representation of users in the network.

Figure 4 – Conceptual model on content selection.



Using the associated study of central nodes and emotional response, I examine whether the content selection is independent of (the presented) network dimensions and exposure of the sender or not (see figure 3). It is expected that the receiver's emotional reaction will be more determinant for content selection (individual’s choices of contents) than social relevance. In this study, social relevance regards the number of adjacent connections and corresponding tie strength of a given node relative to other nodes of the network. The number of adjacent nodes will be found by measuring the value of degree centrality (Freeman 1978). Four hypotheses emerge directly from this.

Hypothesis 1: *there is a relationship between sender's popularity and content selection.*

Hypothesis 2: *surprise response is associated with content selection.*

Hypothesis 3: *surprise response is associated with the quantity of published content by the sender.*

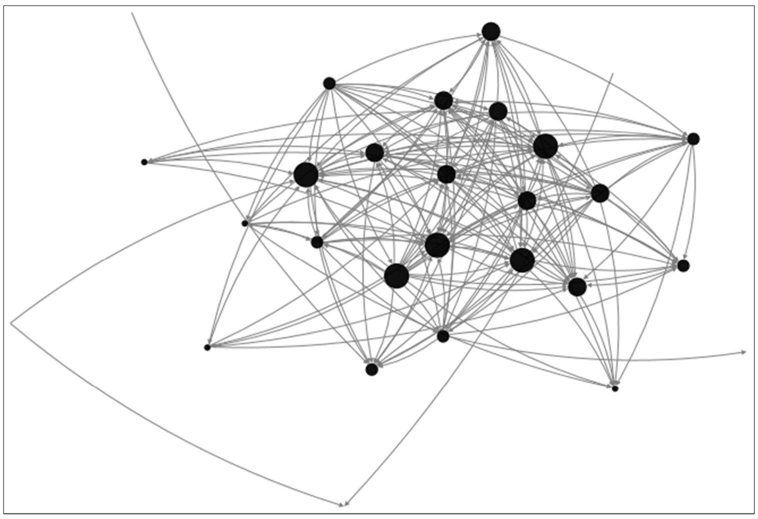
Hypothesis 4: *tie strength is associated to content selection, independently of whether the tie between sender and receiver is a bridge or not.*

7.3.1 Sampling characteristic and procedures

The sample is the same than the one presented in the previous chapters. 56 emotional responses to content selections were validated in the study from 28 participants (16 males, Mean (M) = 19.7 years, Standard Deviation (S.D.) = 1.4 years, 12 females, M = 21.7 years, S.D. = 5.1 years). Similarly, the procedures for data collection were like described in the previous chapter.

The methodology used allowed to produce two different networks. The first network (Figure 4 – Participant network) presents a network of social ties (friendship). With the data on social ties from the network and the information on tie strength from the questionnaire (third stage of the survey – see Appendix A), the entire network of friendship was identified. Degree and betweenness centrality of the senders was measured using the data from this sociograph.

Figure 5 – Participants' network [60].

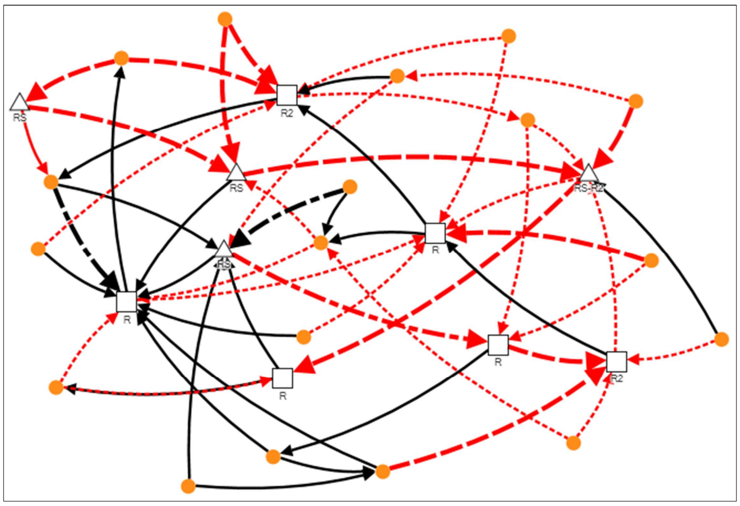


[60] Download from the PFP through the software Node XL.

The second network (Figure 5 – Information flow network) shows ties that represent the individuals' choices of content. Additionally, that will be used to inform us about the emotional reaction to content selection, notably surprise.

Two groups of ties were identified from the Information flow network data. The first group, called "clique", is formed by the participants among whom there is a tie of friendship (either weak or strong). The second group, called "acquaintances", includes the participants with no shared ties between them. This method gives us the possibility to evaluate the number of content selections from a given sender and the corresponding tie strength with the receiver. I could also obtain information about which kind of tie corresponds to a surprise response.

Figure 6 – Information flow's network [61].



Legend:

(→) The arrows mean the direction of whom (receiver) selected a content from whom (sender). It explains the information flow between sender and receiver.

(->) The arrows formed by dots means content selection between sender and a weakly tied receiver.

(→) The arrows formed by lines represent a content selection for a strong tie.

(---->) The arrows formed by dashes means a content selection with surprise response between weak ties (n=12). The results obtained reveal 12 surprise responses.

(-.-.->) The arrows formed by dashes and dots means surprise response too, but in this case there is a strong tie between sender and receiver (n=2).

(□) The square shape (R) means a receiver that was surprised. The number, e.g., R2, means that the participant was surprised by two different content.

(△) The triangles represent a participant that was surprised as receiver and who, as sender, caused surprise with their contents.

(○) The circles mean no surprise.

[61] Configuration obtained with the data validated (56 content selections) using the software Node XL.

Dependent and Independent Variables

The variable defining content selection activity is the main dependent variable: “Nr. of selections”. I expected to identify how receiver’s content selection (Nr. of selections), is influenced by both tie strength with the sender and structural position of the sender in the network. We will observe whether or not the emotional reaction of surprise (given by the independent variable of “Structural holes”) explains the content selection, and if it prevails over the other network dimensions in this regard (“Centrality” and “Tie Strength”). Similarly, the importance of senders’ exposure (“Total of published contents”) through their content published in regard to receiver’s content selection (Nr. of selections), is analyzed. The sender of the content selected is the central node observed.

Relatively to independent variables, four variables were considered.

Contents’ choices describe the participants’ activity in the “Information flow’s network” and was defined by two variables: “Total of published contents” and “Nr. of selections”. Both variables are presented in the tables below under the "Content’s choices", however, “Total of published contents” is the only one as independent variable.

Centrality was measured by degree and betweenness (Freeman, 1979). Degree centrality was computed in both networks. In the “Participants’ network” the popularity of the sender is measured by its degree centrality[62]. In the “Information flow’s network”, the degree indicates the number of receivers who selected the content. Betweenness centrality[63] was computed in the “Participants’ network”. This measure is regularly used as a bridging proxy (Mori et al., 2005; Kratzer & Lettl, 2008). With the values obtained it was verified the relationship between the structural position of the individuals acting as bridges and its values of betweenness.

[62] As reviewed in fifth chapter, degree can be calculated by counting the number of links for each node and, often, it is interpreted as a grade of popularity, prestige, or influence (Knoke & Burt, 1983).

[63] This measure refers to the probability that a ‘communication’ between two individuals takes a particular path, which minimize the number of intermediary nodes, being its length defined as the minimum number of ties linking these two individuals, either directly or indirectly.

Structural holes that connect otherwise disconnected individuals (Burt, 1992) were evaluated through triad-level measurements (Kalish & Robins, 2006). This analysis was based on the identification of triads[64].

Tie strength was based on a construct of four dimensions and dichotomized as weak ties and strong ties, as described in sixth chapter. These ties are described by six variables, which are distributed among two groups: the “clique” group and the “acquaintances” group. These six variables are listed in the item “Tie strength and content selection”, as shown below in the tables 8 and 9.

7.3.2 Results

Descriptive statistics for variables used in the analyses of content selection and publishing describe the values relative to each participant (N = 28).

Table 8 – Descriptive statistics on content selection and publishing.

Participants	N	Minimum	Maximum	Mean	Std. Deviation
Content' choices					
Total of published contents	199	3	22	7.10	4.42
Nr. of selections (first two selections)	56	0	9	2.00	2.37
Clique (group of tied participants with content selected from each other)					
Participants in the clique	160	1	17	5.71	4.13
Strong ties in the clique	70	0	8	2.50	2.26
Tie strength and content selection					
Weak ties from the clique that selected	10	0	2	.357	.558
Strong ties that selected	23	0	7	.821	1.54
Weak ties that selected	33	0	5	1.17	1.56
Strong ties that didn't select	47	0	6	1.67	1.76
Weak ties not from the clique that selected	23	0	4	.821	1.33
Weak ties from the clique that didn't select	71	0	6	2.53	1.87

[64] As mentioned in sixth chapter, a triad is a set of three persons that tend to close through a third person, forming a triadic closure, in which the strength of the ties among individuals plays a determinant role.

Table 8 shows that despite the high number of strong ties in the “clique” group (“Strong ties that selected” and “Strong ties that didn’t select”), the number of strong ties that select contents from someone in the “clique” group is quite small. This seems to reveal the low level of relevance of friendship in decision making for content selection. The Wilcoxon signed-rank test is used to verify this association. On the other hand, the number of “Weak ties not from the clique that selected” (N=23) are much higher than the number of “Weak ties from the clique that selected” (N=10). This seems to reinforce the idea that the level of friendship, even between acquaintances, is not as relevant as distant weak ties on the decision making for content selection. Given the relevance of distant ties as bridges (as verified previously in chapter 6), this seems to be relevant to show the importance of bridging ties – associated to novelty perceived – on content selection.

Table 9 – Pearson correlations.

Variables	(1)	(2)	(3)
Content’ choices			
(1) Total of published contents			--
(2) Nr. of selections (first two selections)	.032		
Centrality measures	--	--	--
Structural holes			
(3) Surprise		.010	
Tie strength and content selection			
Weak ties from the clique that selected		.020	.005
Strong ties that selected		.003	
Weak ties that selected		.011	* X2 (1) = 22,19
Strong ties that didn’t select			
Weak ties not from the clique that selected			.003
Weak ties from the clique that didn’t select	--	--	--

* p < 0.001

The Table 9 shows a positive correlation between the “Nr. of selections” and ties from the “clique” group: weak ties (r = 42.903, p = .020) and strong ties (r = 83.707, p = .003), showing that both acquaintances and friends selected contents. However, no correlation was found between the number of ties of each sender (degree centrality in the “Participants’ network”) and content selection, or in other words, between popularity and content selection.

On the other hand, evidence was found for the relationship between surprise and the “Nr. of selections” ($X2 = 18.563$, $df = 7$, $p < .010$). Furthermore, a strong association was found between surprise and weak ties that selected content ($X2 = 22.193$, $df = 5$, $p < .001$). This includes all weak ties that selected, either from the “clique” group ($r = 10.463$, $p = .005$), or not from the “clique” group ($r = 93.889$, $p = .003$).

The associations were examined between the activities of content selection (“Nr. of selections”) and the independent variables represented by the three network dimensions under study. Table 10 does not list any results for tie strength, degree and betweenness centrality, because no association was found between them and “Nr. of selections”.

Table 10 – Coefficients from regression model predicting surprise.

Predictors	Nr. of selections
Content’ choices	
Total of published contents	.216 (.034)
Structural holes	
Surprise	3.733 (.001)

Hypothesis 1 states that *there is a relationship between sender’s popularity and content selection.* Firstly, by computing Pearson’s correlation no association was found between the number of ties held by each sender (sender's popularity) and the selection of their content. Moreover, applying the multiple linear regression[65] with backward variables selection, I found that the degree centrality values in “Participants’ network” is not associated to the values presented by any variable related to ties from the “clique” group that selected ("Weak ties from the clique that selected" and "Strong ties that selected"). A higher value on degree centrality of a sender does not mean a selection of their contents by a receiver, independently whether their bond is weak or strong, when they belong to the clique that selected. Thus, Hypothesis 1 is not

[65] The assumptions of linear regression were verified.

confirmed and factors underlying to social relevance are not determinant to content selection. Content selection was found not to be associated to the sender's popularity. Individuals make their content choices irrespectively of the kind of relationship they have at friendship level with the sender. As seen in Table 8, the findings show a low level of relevance of friendship on decision making for content selection. Therefore, to have a strong tie with the sender is not predictive of content selection.

To test Hypothesis 2, which posits that the *surprise response is associated with content selection,* multiple linear regression it was applied. It was found that content selection is strongly associated with surprise response (B adjusted = 3.733, $p = .001$). Thus, Hypothesis 2 is confirmed.

To test Hypothesis 3, which states that *surprise response is associated with the quantity of published content by the sender,* it was applied the Fisher's Exact test. The results suggest that there is no association between the activity of publishing contents and becoming more surprised ($p = .433$). Thus, the surprise response is not associated with the contents sender's production, and Hypothesis 3 is not confirmed.

Hypothesis 4 posits that the *tie strength is associated to content selection, independently of whether the tie between sender and receiver is a bridge or not.* Wilcoxon's signed-rank test was applied to analyze whether or not there is a difference between strong and weak ties for content selection (Table 8). The results shown that there is not a significant difference between the number of selections made by strong and weak ties ($Z = - 1.052$, $p = .293$). This seems to indicate that friendship (strong ties) is not prevalent for content selection. Furthermore, by applying multiple linear regressions, no relationship was found between weak or strong ties and the variable "Nr. of selections" (Table 10). The same results were found when surprise was included in the regression model. Thus, content selection is not associated with the strength of the tie. Friendship ties (i.e., strong ties) do not predict content selection, even when this tie is associated with a bridging factor, which confirms Hypothesis 4.

Therefore the results seem to reveal that people make their content selection independently of the tie strength and sender's content production and popularity. Sender exposure does not determine the content choices when the individual is surprised. Once again it is verified that content selection does not obey to social relevance factors, because individuals' choices do not rely on senders identified by high values of centrality measures (i.e., degree centrality, and so, popularity). On the other hand, given that surprise is representative of the perception of novelty, it was found that there is a preference to select contents from central nodes represented by bridges and posted by weak ties, rather than those posted by close friends or associated to centrality measures (Freeman, 1979).

Finally, the coincidence between the assessment of betweenness centrality and bridging was analyzed. Betweenness centrality identifies the brokering position of the participants in the "Participants’ network" and is associated to the network of ties among participants. Bridging associated with novelty perception is identified by the emotional response of surprise in the "Information flow's network" given the content choices. Pearson’s correlation shows that there is not any association between betweenness and bridging. The positions associated to the brokering activity do not coincide with the location of bridges eliciting surprise. Thus, the brokering positions defined by high values of betweenness centrality do not coincide with the positions occupied by senders that elicited the surprise response in receivers.

7.4 Discussion

It is known that bridging nodes are typically located at the periphery (Valente & Fujimoto, 2010), but the broker’s role can also be measured by betweenness centrality and still be independent of degree, which indicates peripheral locations (Haythornthwaite, 1996). Nevertheless, none of these possibilities reveal how content choices are made, because such measures are typically centered in senders’ perspective. This study considered the information flow in a network from the receivers’ viewpoint (rather than the sender's perspective), regarding the two types of central nodes, bridges and the ones defined by centrality measures.

Furthermore, it was analyzed how tie strength may influence the choices of contents by receivers, as well as if the content production of the sender, their exposure in the network, could influence such choices.

The overall results do not confirm hypotheses related to *centrality measures*, but they do confirm the ones related to *bridging factors*. This suggests that only one of the central nodes (i.e., the information flow through network bridges associated to the perception of novelty) matches with the individuals’ content choices. Furthermore, I verified the low level of importance of friendship on decision making for content choices, at least for the contents analyzed in the survey. The contents shared among participants were mostly videos (some of them were songs), photos and online news.

The discussion will be structured into two parts; a) flow of information determined by central nodes associated to high values of centrality measures; b) emotional factors related to central nodes identified by its role as network bridge.

a) Flow of information:

First, I analyze the flow of information in a network considering the viewpoint of centrality measures. Literature on central nodes has been focusing its attention in the benefits associated to network position, either related to the degree centrality (and other derived measures), or related to the brokerage activity. This kind of approach is strongly associated to the role of the individual located in such position, i.e., the position of the sender. Nevertheless, we can extend the understanding about the information flow in the network by considering the receiver's perspective relatively to senders' position, as well as the personal attributes. This may change the assessment made about the importance of a given type of central node. However, a different outlook has been adopted by scholars.

It is correct to say that centrality measures of a central node are not about isolated attributes of individuals, nor are they about their role as a sender of information; rather, they represent the individual's relationship within the network and ability to control the flow of information. From the viewpoint of the number of adjacent nodes, these central nodes are weighted by their social relevance to other nodes, and, thus, are frequently seen as objects, rather than sources, of communication (Knoke & Burt, 1983). It is in such conditions that the benefit (and power) underlying its network position is estimated. However, it is not estimated how that network position, given individual attributes, may benefit other nodes (e.g., giving access to novel information).

On the other hand, those metrics are relative measurements because they compare their elements among each other based on a static structure corresponding to a certain moment in time (Nanda & Kotz, 2008). Of course that, despite the limitation of the metrics, methods and tools to observe such dynamic relationships, seems relevant to advance in studies that comprehend such dynamics in a more holistic way.

b) Emotional factors:

Second, considering the arguments above, it seems that scholars have overlooked how personal attributes (Aral et al., 2009) and, consequently, individual choices may interfere in the information flow. There is a lack of data concerning the individual's role in linking parts. In fact, neither does Granovetter deliberate on how the individuals at each end of the tie participate in the effectiveness of the bridging connections, nor does Burt clarify whether or not the bridging factors are independent of recipients' perception.

I discuss how emotional factors (i.e., surprise) are a better descriptor of content choices than the social relevance factors of nodes.

In Burt's (1992, 2004) concept of bridging it is stated that the differences in interests and unique perspectives of individuals surrounding structural holes creates advantages in the access to information, novelty and the spreading of information (Bakshy & Rosenn, 2012). The individual that spans the structural hole, or the broker that mediates the access to resources by connecting parties or preexisting ties between parties (Katz & Tushman, 1981), transports information on personal attributes and people's social world immersed in the content shared. In this sense, the filtering of information through the network processes creates an interchange of information about people participating in the bridge (Burt, 1992)[66].

This view asserts that it is the network that promotes and legitimates both information and network members, which, from this standpoint, are instrumental in receiving and forwarding such information (Haythornthwaite, 1996). This is a structural outlook that is emphasized by some realms of literature that argue that nodes or groups of nodes of a network can be replaced with no information flow breakdown (Sarr et al., 2012). However, this seems to be an incomplete view when considering the dimension of the psychological characteristics (personal attributes) of the actors in a social network, as the results presented in this work seems to show.

Furthermore, the literature on central nodes usually debates the benefits accessed by the central position occupied, but the overall process behind the structural bridges is not fully characterized, or terminated, with the argument that brokers facilitate the access to novelty (Obstfeld, 2005).

[66] As stated by Burt (1992, p. 14), "the network that filters information coming to you also directs, concentrates, and legitimates information about you going to others".

This study complements the traditional structural view (e.g., Burt, 1992; Valente & Fujimoto, 2010) introducing the surprise response as a proxy of novelty to analyze the emotional reaction to content selection, and presents a different perspective on how social network dimensions may influence content choices of the Web users. It was also shown that that content selection in a social network environment is more dependent of receiver's emotional reaction than from factors associated with node's social relevance – characterized in this work by popularity of the sender and tie strength between sender and receiver.

In summary, though popularity and friendship suggest that a network's central positions show nearness, these two dimensions are not associated with receivers' content selection. This association can also not be made with all kinds of bridges (structural holes), but only with those related to the receivers' stimulus of surprise. Therefore, the overall results indicate that the network dimensions of centrality (degree and betweenness) and tie strength (i.e., friendship) are less relevant to content choices than has been assumed (considering the relevance attributed to those network dimensions on providing social data and solutions to digital media systems). Instead, structural holes spanned by weak ties reveal a strong relationship with receivers' choice of contents. Particularly, by applying surprise as proxy of novelty perceived, the relevance of the emotional reaction in the content choices is made clear.

7.5 Conclusion

This study generalizes Burt's (1992) assertion about bridges, highlights its relevance as a central node, and the importance of novelty perception to validate bridging factors (i.e., non-redundant structural holes), and study the behavior of content selection by receivers of information.

Bridging nodes present valuable arguments as central nodes, either by the uniqueness of the information flow brokered or by their association with the emotion of surprise. This allows the creation of content personalization rich in new perspectives for the receiver, and offers a good representation of people's behavior when they select contents. For specific concerns, they are a valuable alternative to central nodes identified by centrality measures.

More studies are required concerning the bias set forth by the information flow centrality measurements, which is centered in the number of ties, as proposed in the original measures (Freeman, 1979). This emphasis on centrality may have weakened the development of other measures for structural positions (Valente & Fujimoto, 2010), in particular, understanding the role of users' psychological characteristics in social networks.

The development of a different approach, e.g., considering the prediction of surprise, can have significant applications in digital media systems, such as in recommendation systems and search engines. Considering the current demand for social data, scholars may be encouraged to extend the study of emotions elicited on social networks, notably from the perspective of the perception of novelty.

This research may have limitations, given the sample used. An extension to these results could be found by analyzing how the cognitive distance between receivers and senders may justify the stimulus of surprise. In doing so, analyzing which factors could justify similarities at an emotional level that could determine an optimal cognitive distance for the perception of novelty becomes equally significant. The next chapter of this dissertation presents an approach to the analysis of these questions.

CHAPTER 8

PERSONAL ATTRIBUTES AND BRIDGING TO DEFINE COGNITIVE DISTANCE: PREDICTING SURPRISE

8.1 Overview

Our network of contacts and level or interaction with which we usually do not have a frequent contact is growing fast. This is raising the importance of communication among people bound by a weak tie and so, the need to understand the data behind the connections between people socially distant, through which novelty can be exchanged. However, little attention has been given to the implications that personal attributes may have in this process, notably, in the information flow from the standpoint of an individual’s emotional reaction when the information is accessed. Thus, I test which personal attributes (i.e., homophily, preferences of music genre and emotional reaction to music genres, and political views) and bridging factors represent the optimal cognitive distance that is associated with the perception of novelty. Here, surprise is applied and justified as a proxy of novelty perceived by receivers. Findings show that dissimilarity rather than similarity compose the cognitive distance that explains the surprise response, jointly with bridging factors. These dimensions are relevant to design personalized recommendation based in novelty.

8.2 Introduction

Interactive media like online social networks have been scaling our access to a larger number of people, which mainly consists of acquaintances instead of people with which we have a frequent contact. In this context, weak ties are becoming more influential than strong ties on behavior or opinions that people choose to adopt (Bakshy & Rosenn, 2012). This fact creates a totally new kind of output based on sharing views between Web users, which can benefit from new insights and novelty, as the theories on network bridges show (Granovetter, 1973; Burt, 1992). Furthermore, there is an association between surprise and bridging factors (i.e., weak ties and non-redundant structural holes), connecting senders and receivers of novel information, as we have seen in the sixth chapter. This highlights the need to know more about how to collect data on the different possibilities of users' interaction beyond the ones based on the homophily dynamics, adjacent connections, strong ties, or characterized by centrality measures. However, little attention has been given to the implications that personal attributes may have in the information flow from the standpoint of an individual's emotional reaction when the information is accessed. This knowledge is important, in order to characterize the processes of interaction between people socially distant (i.e., connected by weak ties or by means of structural holes). Further, this data is easy to get from social networks. Thus, in this chapter I examine the role of similarities/ dissimilarities between sender and receiver when surprise is elicited. In order to do that a threefold approach is proposed.

Firstly, I applied the concept of optimal cognitive distance (Nooteboom, 1992; 2005). This conceptualization states the importance of differences in cognition (cognitive distance) between individuals and the trade-off between a higher novelty value and a mutual understanding. Where this distance is too large, it may preclude mutual understanding and then the information received will not be perceived as novel. While if it is too short, this means that there is too much familiarity in the information shared (Nooteboom et al., 2007) and, thus, no surprise involved, given the absence of novelty (Barto et al., 2013). Distance in Nooteboom model is explained by means of the existing dissimilarity between partners and by the contribution to the creation of new knowledge and novelty (Nooteboom, 2000, 2007).

Secondly, I propose a way of solving the issues caused by the absence of a direct measurement of cognitive distance. Wuyts et al. (2005), who tested the optimal cognitive distance hypothesis in the perspective of finding the technological and organizational differences between partners of pharmaceutical firms, identified that as a major limitation[67]. To surpass this constraint, although in a different context of application, this study proposes the use of personal attributes and test network dimensions (i.e., bridging factors) to define such a distance between sender and receiver. I propose a range of personal attributes to identify the optimal cognitive distance underlying the perception of novelty, by means of detecting the surprise elicited when the receiver selects a content of a sender. The range of personal attributes

[67] The authors assumed that the more that pharmaceutical firms cooperate with the same partners over time in their agreement portfolio, the lower will be the average cognitive distance with their partners. They argue that the assumption is consistent with an earlier finding that cognitive distance decreases as interaction is more frequent (e.g., (Lewicki and Bunker, 1996).

comprises the dimensions of status homophily (McPherson et al., 2001), attitudinal similarity (McCroskey et al., 1975, 2006), political views (Lin & Ensel, 1981, Fond & Neville, 2010), preferences of music genre (Rentfrow & Gosling, 2003), and emotional reaction to music genres.

Thirdly, the assumptions on bridging proposed by Granovetter (1973) and Burt (1992) are tested. It is well established that the information flow crossing a bridge (Granovetter, 1973; Burt, 1992) in a social network is strongly determined by the level of novelty that it carries to a receiver.[68]

Therefore, the aim is to identify which personal attributes and bridging factors jointly characterize the optimal cognitive distance underlying the perception of novelty, i.e., surprise.

That goal has however an important implication, that is the need to combine the cognitive view with the social structural view or, in other words, to analyze the association between personal attributes and the bridges outputs.

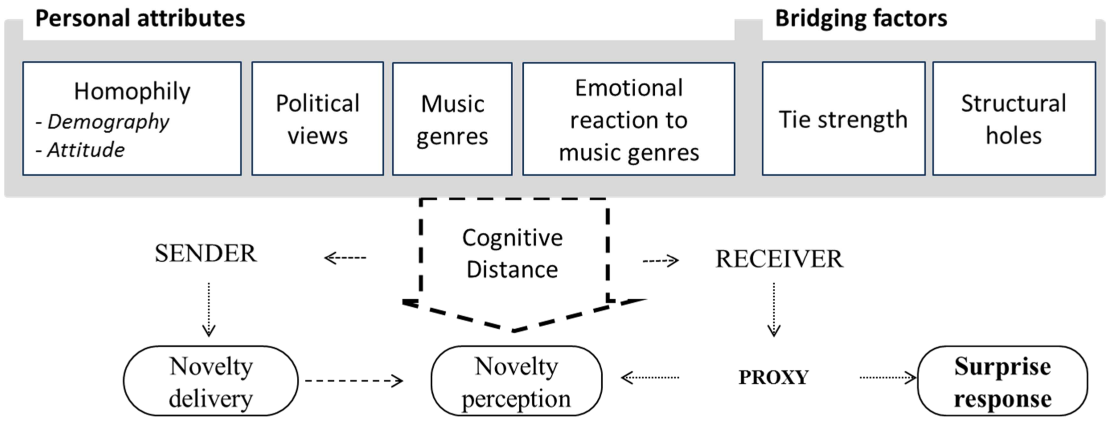
Those goals also give the possibility of exploring an alternative approach to social network analysis, notably in the understanding of the delivery of novelty through network bridges, as well as in the use of these social data in Web applications, like social-based recommender systems. That line of reasoning may introduce a more detailed knowledge in which dimensions characterize the interaction between two socially distant people in a network when a specific emotion is elicited, i.e., surprise. Furthermore, the interplay between bridging and emotional reaction may show the way towards the next generation of social networking for digital media systems and a new approach for scholars in the field of social network analysis.

8.3 Cognitive distance and hypothesis

In this chapter I examine in a social network environment which personal attributes and network dimensions (i.e., bridging factors) are associated with the surprise response when a receiver selects contents. As Figure 6 shows, in such conditions there is a cognitive distance between sender and receiver into which surprise is elicited.

[68] In this study, similarly to what was done in the previous chapters, surprise (e.g., Teigen & Keren, 2003) is used as a proxy of novelty (Stiensmeier et al., 1995; Strange et al., 2005).

Figure 7 – Conceptual model on cognitive distance.



Nooteboom's optimal cognitive distance hypothesis was already tested by several scholars (e.g., Wuyts et al. (2005) and Nooteboom et al. (2005)) in order to explain the inverse U-shaped relation between novelty and cognitive distance. The optimal level is found here at the middle point between the very low and very high levels of cognitive distance. This is related with the tradeoff between opportunity and challenge in processes of learning and innovation, in interaction between firms (Nooteboom, 1992; Nooteboom, et al., 2005, Wuyts et al., 2005).[69] The opportunity is related to diversity, where the novelty value of a relation increases with cognitive distance. The challenge lies in finding partners at sufficient cognitive distance to tell something new, but not so distant as to preclude mutual understanding. In this sense, Nooteboom (2005) posits that with more knowledge one needs larger cognitive distances to find novelty. In a similar vein, Gilsing et al. (2008) state that, cognitive distance refers to the extent that, organizations differ in their technological knowledge and expertise. Here, the authors consider the role of cognitive distance among organizations forming an alliance network.

Regarding this study, the concept of optimal cognitive distance was adopted with the aim of framing a possible range of dimensions based on similarities (or dissimilarities) and network factors that justify the surprise response when a content

[69] The results found were tested in 994 alliances in several industries, in the period 1986-1996, by Nooteboom et al., and on interfirm agreements between pharmaceutical companies and biotech companies, as well as on interfirm agreements in ICT industries, by Wuyts et al..

shared by a sender is selected by the receiver. Two reasons justify this option: a) The inverse U-shaped relation between novelty and cognitive distance helps to frame theoretically the approach of this study; b) The adequacy of using surprise instead of novelty, it is adequate and justified by the proxy between novelty and surprise introduced in previous chapters.

In this sense, six hypotheses will be explored. The listed hypotheses incorporate two opposite views to justify surprise response. One is based in the similarities of personal attributes between sender and receiver. The other is based on dissimilarities. Each hypothesis on similarity and dissimilarity is also tested with bridging factors. I do not list here mixed hypotheses on similarity and dissimilarity, e.g., similar in music but dissimilar in political views, and vice versa. However, the results of these tests are debated in discussion section.

Hypothesis 1: *Surprise is elicited when sender and receiver share dimensions of status and attitude homophily and have similar interests in music and political views (Homophilous Hypothesis).*

Hypothesis 2: *Surprise is elicited when sender and receiver are dissimilar (Dissimilar Hypothesis).*

Hypothesis 3: *Surprise is elicited when sender and receiver are bridged by a weak tie, share dimensions of status and attitude homophily and have similar interests in music and political views (Homophilous and weak ties Hypothesis).*

Hypothesis 4: *Surprise is elicited when sender and receiver bridged by a weak tie are dissimilar (Dissimilar and weak ties Hypothesis).*

Hypothesis 5: *Surprise is elicited when sender and receiver are bridged by non-redundant structural holes, share dimensions of status and attitude homophily and have similar interests in music and political views (Homophilous and structural holes Hypothesis).*

Hypothesis 6: *Surprise is elicited when sender and receiver bridged by non-redundant structural holes are dissimilar (Dissimilar and structural holes Hypothesis).*

8.3.1 Setting

Sample and procedures for data collection were the same as the ones presented in previous chapters. 56 emotional responses to content selections were validated in the study from 28 participants (16 males, Mean (M) = 19.7 years, Standard Deviation (S.D.) = 1.4 years, 12 females, M = 21.7 years, S.D. = 5.1 years).

Dependent and Independent Variables

The dependent variable in this study is the *surprise* perceived by participants when they selected shared contents. As previously discussed, surprise is an accurate proxy to study the receivers' novelty perception while receiving information through a bridge. Here, two bridging factors are observed separately. One concerning Granovetter's (1973) proposal, based on the weak ties. The other, based on Burt's (1992) theory of structural holes.

Relatively to independent variables, two groups were considered. The first group includes the network factors characterizing bridging assumptions. The two bridging factors analyzed are relative to the variables of tie strength (Granovetter, 1973) and non-redundancy (Burt, 1992). The procedures for measuring tie strength and structural holes were described in previous chapters.

The second group of independent variables refers to personal attributes. It includes five variables: a) socio-demographic dimensions; b) attitudinal similarity; c) political views; d) preferences of music genre; e) emotional reaction to music genres.

Socio-demographic dimensions: Each participant characterized their own dimensions on status homophily. For the study it was considered the dimensions of economic factors, gender, ethnicity and religion (McPherson et al., 2001). Given the sample homogeneity, we withdrew the dimensions of age and educational level. Each participant characterized their own socio-demographic dimensions by answering the online questionnaire. Status homophily data was collected individually. For normalization of such data it was estimated the euclidean distance[70] between receptor and source for each dimension of status homophily.

Attitudinal similarity: The Perceived Homophily Measures (PHM) of McCroskey et al.'s (1975, 2006) was adopted to evaluate the attitudinal homophily.

[70] Euclidean distance gives a measure of dissimilarity between two variables.

McCroskey et al. (2006) model allow the study of variables such as *Attitude* and Economic *factors*, regarding the perception of others. This model fits well the approach to the Attitude[71] study.

Political views: Participants were asked to specify their political affiliation, or political inclinations. Five options were listed: Conservative; Moderate; Liberal; Independent; Other. To compute the result the variable was dichotomized.

Music genres preferences: preference of music genre and emotional reaction to music genres were based on the use of the dimensions studied by Rentfrow & Gosling (2003) about musical preferences[72]. Participants were asked about their preferences in musical genres and classified different types of music by selecting a value ranging from 1 (Very negative) to 10 (Very positive). Four categories of music genres were presented, such as “Reflective and Complex (Blues, Folk, Classical, Jazz)”, “Intense and Rebellious (Alternative, Heavy metal, Rock)”, “Upbeat and Conventional (Country, Religious, Pop)”, “Energetic and Rhythmic (Funk, Hip-Hop, Soul, Electronica)”.

Participants were also asked about their emotional reaction when they listen to a particular type of music, based on the same four categories of music genres listed above. Participants classified each category with one specific emotion out of a list of ten. These emotions were based on the DES scale of Izard (1991). The variables “preference of music genre”, and “emotional reaction to music genres” were both dichotomized.

[71] Participants answered a set of six questions based on a five-point semantic differential scale. Scale items included descriptors such as “The participant that stimulated the emotion of ‘surprise’ on me”: ‘Behaves like me (e.g., in public, among friends)’. In order to dichotomize this variable, firstly a value from 1 to 5 was attributed for each item of the scale (e.g., 1 – “Strongly disagree”; 5 – “Strongly agree”). Secondly, the mean and the standard deviation were estimated. Then, the lowest value of the scale (A – lower border) was obtained by subtracting the value of the standard deviation from the mean value. By adding the standard deviation value to the mean value we found the other end of the scale (B – higher border). All the values lower than or equal to A and equal to or greater than B were considered in the extremes. We dichotomized the variable by coding the extremes (A and B) with 0 and 1 (between A and B).

[72] The authors used a set of music’ genres already studied (Reflective and Complex, Intense and Rebellious, Upbeat and Conventional, and Energetic and Rhythmic) to identify, or predict, traits of personality according to a wide array of personality dimensions (e.g., openness), self-views (e.g., political orientation), and cognitive abilities (e.g., verbal IQ). The authors’ claim is that music preferences are partially determined by personality, self-views, and cognitive abilities. For the questionnaire, we adopted the framework used by Rentfrow & Gosling (2003).

8.3.2 Results

The study aims to draw conclusions about which personal attributes and bridging factor predicts surprise. By means of the dimensions mentioned, it was intended to characterize the optimal cognitive distance between sender and receiver underlying the perception of novelty. This means, when the surprise is elicited.

Firstly, it was analyzed which personal attributes have a relationship with the perception of novelty through the proxy of surprise. Secondly, it was analyzed the association between each bridging factor and surprise and on how this structural property interferes in the relationship between personal attributes and the perception of novelty.

8.3.2.1 Personal attributes and surprise

The first step was to establish the association between surprise and the independent variables describing personal attributes. Pearson's correlation results (see table 11) suggests that gender (X2 = 4.691, p = .030) and attitudinal similarities (X2 = 4.058, p = .044) are the only variables associated with surprise. This means that in all content selections related to surprise, 74% correspond to different genders between receivers and senders and that in 64.3%, receivers consider themselves similar to the source of information.

Table 11 – Pearson's correlations.

Variables	(1)	(2)	(3)
Economic factors	--	--	--
Gender	--	--	.030
Ethnicity	.017	* 12.355 (.001)	--
Religion	0.48	.012	--
Attitudinal similarity			.044
(1) Tie strength			.019
(2) Redundancy	* 20.541 (.001)		.005
(3) Surprise			

* p < 0.001

To study the correlations between the variables it was performed logistic regressions.

Internal consistency reliability was established by Cronbach alpha values. It was computed such coefficients for attitudinal homophily (.67) and tie strength (.97).

Table 12 – Coefficients from regression model predicting surprise.

Predictors	Tie Strength	Redundancy	Surprise
Personal attributes			
Economic factors			--
Gender			4.062 (.037)
Ethnicity	-.224 (.021)	-.096 (.001)	--
Religion			--
Attitudinal similarity			-.243 (.054)
Interactions			
Gender (& tie strength)			4.379 (.037)
Gender (& redundancy)			--
Tie strength (& gender)			-.394 (.029)
Redundancy (& gender)			--

Logistic regression was used to explore the influence of the variables of the socio demographic variables (religion, ethnicity, economic factors and gender) on the variable of attitude. None of the tested variables presented an association with attitude.

The association between tie strength and the variables of ethnicity, gender, economic factors and religion, as well as attitude was also analyzed. Applying logistic regression, the results showed that only ethnicity was significantly related with the tie strength. The odds of having a strong tie decreased when the sender and the receiver had different ethnicities (odds ratio = .224, p = .021, 95% CI [.062, .801]).

To understand whether any of these independent variables could be a predictor of surprise, logistic regression was computed (see table 12). The results suggest that either attitudinal similarities (odds =.243, p = .054, 90% CI: [.073, .813]), or gender (odds ratio = 4.062, p = .037, 95% CI [1.089, 15.150]) hold a significant relationship with surprise response. When the sender and the receiver have different genders, the odds of having a surprise response increases. The same happens regarding the attitudinal similarities between these actors. When there are attitudinal similarities, the odds of having a surprise response increase.

8.3.2.2 Personal attributes and bridging factors of surprise

Next, it was mainly examined the associations between each bridging factor and personal attributes with surprise. Firstly, it was found that there is strong evidence of a relationship between surprise and tie strength (X2 = 5.534, p = .019), where 36% are weak ties and 9% are strong ones, and 85.7% of the weak ties spanning non-redundant structural holes (N = 11) are associated with the stimulus of surprise (see table 11), as well as between surprise and redundancy (X2 = 7.754, p = .005), where 78.6% of these connections are established over a structural hole linking non-redundant peers of receivers and senders of information.

Secondly, logistic regression (table 12) was computed separately for each bridging factor jointly with personal attributes as independent variables, and with surprise as a dependent variable. Including personal attributes and tie strength in the regression model, both gender (odds ratio adjusted = 4.379, p = 0.037, 95% CI [1.090, 17.587]) and tie strength (odds ratio adjusted = 0.394, p = 0.029, 95% CI [0.171, 0.908]) were significantly related to surprise response. The odds of having a surprise response decreased with strong ties. Nonetheless, when the regression model included gender and redundancy, no significant relationship with surprise response were found.

8.4 Discussion

This study tested a range of personal attributes to find which one is associated with the surprise response when a content delivered by a bridge is perceived by a receiver as novelty. This means identifying the optimal cognitive distance measured by personal attributes that jointly with bridging factors predict the surprise.

This study consisted of two levels of analysis. First, the association between personal attributes and surprise was examined. It was found that only two dimensions of homophily, i.e., gender, more specifically gender differences, and the attitudinal similarity were associated with surprise response. The regression model showed that differences in gender and similarities in attitudinal behavior, analyzed separately, predict surprise. These findings suggest that these two personal characteristics make up the cognitive distance that explains the surprise response.

Surprise is elicited when the cognitive distance between sender and receiver is not too short, and nor is it too great. If the distance is too short, the familiarity of the information will prevent any surprise, and if it is too great, it may preclude mutual understanding to benefit from the opportunity of a novelty perception. It is in between cognitive borders that surprise occurs. Besides that, other emotions are elicited, but which are not relevant for the bridging effect.

It is assumed that cognitive similarities contribute to improving the perception of the message (Roger & Bhowmik, 1970), and that similarity induces homophily (McPherson et al., 2001), as well as to establishing links of trustworthiness, and generate better acceptance of recommendations between pairs (Arazy et al., 2010). Nonetheless, the results suggest that what reinforces the level of communication through the emotion of surprise is the gender difference. Thus, the factor of cognitive agreement is based on heterophily (dissimilarities) rather than on, only, homophily. Here, the optimal cognitive distance associated with the surprise elicited is due to a mix of heterophilous and homophilous factors.

In fact, H1 (*Homophilous Hypothesis*) was confirmed, but only for attitude similarity. H2 (*Dissimilar Hypothesis*) was also confirmed due to the differences in gender. This means that surprise is more probable to occur between two individuals that share information online if they have similar attitudes, or are from different genders. Furthermore, the tests did not show any significant relationship between surprise and similarities or dissimilarities in music preferences and emotional reaction to music genres. The same results were obtained for political views. Mixed hypotheses were also tested, e.g., similar in music but dissimilar in political views, and vice versa, with the aim to define the best possibilities based in personal attributes. Nevertheless, I did not find any valid combination.

Thus, if the model does not change its predictors, this can be taken as evidence that these attributes (i.e., economic factors, religion, music genres and corresponding emotional reaction, and political views) included in the model, do not mediate the relationship between bridging factors and surprise, nor are they predictors of surprise when tested in isolation in the model.

This raises the question of how the relationship between communication agreement and differences in gender, between two actors, challenge the conventional assumptions about homophily (McPherson et al., 2001). Several scholars report that gender homophily is an inductor of tie creation. In this vein of research, van Duijn et al., (2003) and Leenders (1997) posit that gender homophily justifies the formation of friendship ties. However, this is not true in cases where the strength of the tie is not strong, i.e., in “friendly” or “neutral” relations. More recently, scholars have noticed that absolute similarity in individuals’ attributes may not characterize social connections. In some instances, individuals may try to find a balance between similarity in some dimensions and differentiation, or heterophily, in others (Rivera et al., 2010). A similar idea was supported by Blau (1974, p. 622), who stated that: “It may ultimately be an oversimplification to refer to a relationship as homophilous or heterophilous, as few individuals do not differ in at least some dimensions and match in at least a few others.”

Second, the delivering of novelty was tested by applying the bridging factors of Granovetter (1973) and Burt (1992). Both conditions were tested separately. To test the perception of novelty by receivers, surprise was applied as a proxy of novelty and computed as a dependent variable. Surprise represents the emotion related to the appraisal of the information delivered by a bridge, when novelty is perceived. Thus, to be surprised in this context means to perceive novelty.

Novelty “determines whether there is a change in the pattern of external or internal stimulation, particularly when a novel event occurred or is to be expected” (Leventhal & Scherer, 1987, p. 15); and an observation is novel when its representation is not found, or is not similar to another one stored in memory (Barto et al., 2013). Consequently, three reasons seem to justify that the perception of novelty is conceived in a framework of communication: a) there is a process of communication because the receiver interprets the information received, as is shown by the emotion elicited; b) the information when accessed by the receiver was already interpreted by the sender; c) the surprise response that is characterized by an optimal cognitive distance between individuals, explains a mutual understanding and interest in the content shared. Hence, it seems correct to assume that in these circumstances there is an effective communication between sender and receiver based on cognitive similarities that are not fully explained by endogenous effects like homophily[73], similarities based on music and political interests, or even through structural factors in isolation.

When analyzing the relationships described above between the bridging factors of Granovetter (1973) and the five variables included in the personal attributes, only gender remains significantly associated with surprise. Moreover, by considering Burt’s (1992) assumptions to configure the bridging factors, no personal attributes, including gender and attitude, stay in the regression model. Therefore, only the bridging factors based on Granovetter’s (1973) assumptions match the hypotheses listed. It confirms H4

[73] As debated in Roger & Bhowmik (1970), or by McPherson et al. (2001).

(*Dissimilar and weak ties Hypothesis*), but hypothesis H3 (*Homophilous and weak ties Hypothesis*), H5 (*Homophilous and structural holes Hypothesis*) and H6 (*Dissimilar and structural holes Hypothesis*) are not confirmed.

In sum, the delivery of novelty in a social network associated to the surprise response can be depicted by an optimal cognitive distance between sender and receiver. This distance can be defined by gender differences and a structural position defined by weak ties acting as network bridges.

On the other hand, the data do not show differences in people's attitude as a function of socio demographic variables. However, attitude similarities are a relevant factor for surprise response, as well as gender differences. This seems to emphasize the fact that the usual variables used to describe homophily behavior, or its effects, may not be sufficient when the actions under observation are information sharing. This means, when the emotional response is a relevant factor to drive such behaviour.

Regarding the literature on *affect* (which refers to a range of feeling states that includes different emotions), it is reported that affect can determine the network structure rather than the other way around (Totterdell et al., 2004). Accordingly, even regarding it in a very simplified way, it seems adequate to argue that bridges can be seen as enablers that approach people with mutual interest in similar topics. Considering, then, the accurate network factors, i.e., bridging as a structural facilitator, and the cognitive distance between recommender (sender) and user target (receiver of the recommendation) as suitable for accommodating surprise response, the benefits would be twofold. First, it will be able to deliver novelty to recommendees. Second, it will be a potential predictor of tie formation, or on the strengthening of ties. In accordance to the latter, these results seem to open up the issue of how the formation of ties is established across social networks. I discuss this issue in the following.

Scholars have been debating extensively the mechanisms of network evolution that lead to creation and break of ties. Nevertheless, in these discussions the role of emotional response has been disregarded, particularly in the case of surprise. There is still no study on how cognitive distances may influence information sharing, which may overlap or complement the adjacency factors and assortative mixing (Goodreau et al., 2009; Rivera et al., 2010) that justify the formation of ties based on individual attributes.

In that sense, the approach proposed seems to challenge the idea which can be found in published literature that bridging structural conditions are opposed to transitivity and, so, to tie formation that is associated to such endogenous effects (Schafer, 2011). This means that from a structural standpoint, Granovetter’s (1973) hypotheses of “The Strong Triadic Closure Property”, also tested by Shi et al. (2013), contradicts the conditions for bridging formation. Both are accurate. However, by considering the bridges as potential inductors of tie formation (i.e., a structural facilitator), this assumption deliberates on the structural conditions of networks. This includes the cognitive conditions that bring to this discussion the actors’ personal attributes. Thus, bridges can be seen as a mechanism of tie formation, when analysed together with factors of cognitive distance. Moreover, this proposal does not disagree with Granovetter's hypothesis.

A distinct but related body of literature considers that the spreading of information in a social system is content dependent and that it assumes different behaviours in different networks (Holme & Ghoshal, 2008). It seems to reflect the fact that people react to contents differently depending on their emotional interaction with that content. Indirectly, this reaction seems to mimic the way the receiver perceives the sender in the topic exchanged.

As a result of that, I speculate that people shape their networks (not the other way around) depending of the perception on others through the contents shared. To justify this assertion, I argue that the contents are like a proxy that interfaces the emotional and affective contact between sender and receiver. This is a view that highlights the idea that endogenous properties like homophily (McPherson et al., 2001; Aral et al., 2009) cannot fully explain the interactions in a social network, neither the structural position when seen in isolation. This reinforces the argument about the use of social data that considers an optimal cognitive distance between sender and receiver, to counteract the social echo chamber effect, instead of social data based on adjacent connections and similarities.

Furthermore, although the role of psychological attributes in the social network analysis has been downplayed (Crosier et al., 2012), the present work shows its relevance and how a more attentive view of them may extend the understanding of social networks. Additionally, considering the psychological attributes in this study, new light will the shed on the assumption that a network of nodes, notably when they represent people, is more than a homogenous set manipulated by homophily, social influence or structural conditions alone. They are all this and the emotions elicited, at least. Thus, beyond the structural properties, social networks are configured by individuals' behaviour and their attributes. This reflects their activities, interests, opinions and emotions when they interact with content and, so, directly or indirectly, with other people.

8.5 Conclusion

The tests presented showed associations between dissimilarity on gender and weak ties as bridging factors that predict surprise. The approach used shows some promising potential when diverting the study of bridging into a new direction, such as towards social media systems. Moreover, it may direct us towards the next-generation of social networking by suggesting on how to seek for proxies that can be used to predict the delivery of novelty through the information flow, or in a social network-based recommender system (by applying social network data to compute recommendations).

In fact, as reviewed in previous chapters, the emergence of online social networks and the access to its data sparked the rise of social network-based recommender systems. This new approach to online recommendation is based on information provided by users' behavior, social ties and similarities, in order to improve personalized recommendation. However, as already debated in this dissertation, the use of these kind of social data also constrains the reach of the recommendation system as such recommendations may become very similar and, thus, less attractive for the user. The introduction of novelty through the data provided is then very important. The present work shows which network dimensions and users attributes should be considered to design such kind of recommendation.

This work is not without limitations. First, although individuals' views on politics were measured, this was based on a single question. The process undertaken was accurate, but it was not possible to control respondents' differences on socio-demographic and cultural background to avoid different interpretations when they needed to classify political choices.

In subsequent work, it would be relevant to analyze other proxies which can represent the concept of cognitive distance, such as by testing personality similarities using the framework of the Big Five personality traits (Gosling et al., 2004; Back et al., 2010). Other approaches can also contribute to better understanding the role of personal attributes and the way they interplay with structural and relational properties in the bridging effects.

CHAPTER 9

DISSERTATION CONCLUSIONS

The purpose of this dissertation is to examine the challenges associated to personalization of Web-based services, notably by recommender systems, and find a solution for the problem of Social Echo Chamber Effect. The challenges associated with the increasing of web data and the possibilities opened by new uses of social data offer new research lines that I tried to assimilate.

Personalized recommendation based in social data from social networks has been pointed out as a good solution to improve performance and solve persistent problems in these systems. However, as discussed in this dissertation, the use of social data based on relationships set by endogenous effects (i.e., homophily) and friendship or social proximity (i.e., strong ties) creates a new problem in recommender systems which I have named the Social Echo Chamber Effect problem. This term seeks to represent the cause and effect related to the use of social data aimed to improve the performance of personalized recommendation. What this term attempts to explain is different from other ones that also describe problems related to personalization, e.g., “Echo Chamber” (Sunstein, 2009), which explains that people naturally seek those who agree with them, or “Filter Bubble” (Graells-Garrido et al. 2013), which draws our attention to the fact that the Web algorithms prevent people from being exposed to viewpoints different from their own, as discussed in third chapter.

I argue that, the Social Echo Chamber Effect traps people inside social bubbles of information. This is due to the lack of diversity in users’ viewpoints (Vargas & Castells, 2011) that are clustered by endogenous properties and, thus, exposed to the lack of novelty in information delivered (Golder & Yardi, 2010) and shared among them.

In order to find a solution to this problem, I have examined an alternative use of social data, with the aim of delivering novelty to the receiver. With this in mind, I developed an empirical work in the field of Social Network Analysis (SNA), and applied knowledge from neuroscience and psychology on novelty perception and surprise response to support the experimental framework. I have found the need to extend the current methodologies to deal simultaneously with content and the users' cognitive reaction.

On the other hand, Web 2.0 technologies have created tools that make Web-users active participants in social networks that they can now also create and operate by themselves. Trust and spatial proximity, associated with specific incentives like friendship, appreciation, knowledge sharing, democratic participation, financial support, or collective creation (Lai & Turban, 2008), have become the design focus of these systems. Moreover, because this Web of social links is more organized around the users rather than around content, more information on users’ interests and habits has become accessible for computation. In fact, as it was argued in third chapter, the Web is emulating human narratives. This can be found in the implicit information contained in the social links and in the content that is interpreted and shared. As a consequence of this, "meaning", which used to be private, is now mutual and shared with the receiver through the information delivered by the sender, i.e., everyone can now go deeper inside the thinking of others through the information shared.

With this understanding, and the boom of online social networks, the activities of sharing common issues and interests that came to be viewed as the reward of the whole system, also became an advantage for other Web-applications, notably for personalization. Consequently, factors related to friendship (Granovetter, 1973) and homophily (McPherson et al., 2001), associated with the growth of knowledge about users’ individual characteristics, have become key-references to define borders of information. However, when Web-based applications use these naturalistic behavioral data (Boyd, 2007), to create a representation of their users and their networks, these data only mirror social relationships determined (and confined) by social organizing principles based on homophily. This means that the dimension that includes the psychological characteristics of the users is missing, and consequently significant information about individual attributes.

Hence, when these data are used to improve personalization, they are in fact transporting into the recommendation the information from the set of people that share the same echo chambers. Consequently, this kind of personalization is strongly related to the concept of "Social Echo Chamber Effect", as I have stated.

In this sense, it would seem that, once again, the development of a Web technology is not looking carefully enough at the cognitive factors that can limit its success, at least from a user perspective. The reduction of quality of personalization services seems to be related to cognitive factors rather than to technological factors.[74] Thus, there is a technological limitation that is only detectable if the researchers and developers, notably in computer science, are aware of this kind of knowledge, and reach an understanding (and agreement) on the impact that cognitive factors may have on technological development. As a consequence, instead of gaining facilitated access to information, through media, people end up merely spinning around inside their social worlds.

There are uncountable drawbacks related to this restrictive reality. In this sense, we highlight three motives that have a negative impact in social interaction and in the individual behavior, when people access online information based in the current solutions of personalization, which justify the need for alternative solutions. First, the echo chamber is conducive to increased conformity and less diversity. Accordingly, people lose the stimuli to ask new questions, which may reduce learning and creativity.

Second, less novelty is associated to less surprise, which means less richness in the construction of meaning. This fact may reduce the ability to interpret the surrounding reality exploring different perspectives.

Third, less diversity in the viewpoints generated among users, means reduced quality in the services provided by recommender systems, and so, a lower level of satisfaction for these users.

Despite the drawbacks associated to the social data, listed above, it does not follow that social data should be avoided to improve personalization or other types of Web-applications.

With these considerations in mind, what does the present dissertation contribute to our knowledge about how to use social data and avoid the Social Echo Chamber Effect?

[74] A similar hypothesis was argued in the context of the lack of success of the Semantic Web proposal, as well as of other automated services sustained at the semantic level. The reasons detailed in the third chapter for this are related to the different boundaries of the meanings of words and linguistic expressions that vary from person to person. The simple fact that what is expressed in words does not mean the same to everyone, may drastically reduce the opportunities for convergence in these automatic services.

In order to answer to this question, three studies were developed in order to analyze social interactions from the perspective of the receiver of information (who makes the content choices). The aim was to find which dimensions are behind social data, i.e., structural factors and personal attributes, that contribute to the perception of novelty, with the purpose of provide a new kind of data source for personalization. In particular, improve the performance of social network-based recommender systems.

Three empirical studies presented in Chapters 6, 7 and 8, respectively, consider in tandem the study of social ties and similarities among a population of students (participants in the empirical work undertaken) and the emotional response to content selection in a social network environment. This option had a twofold aim: a) To conceive an appropriate methodology to study the problem presented by the Social Echo Chamber Effect; b) Extend the current approaches on SNA, by researching the role of emotions, in particular surprise, as well as its relationship with personal attributes, such as dimensions of status homophily (McPherson et al., 2001), attitudinal similarity (McCroskey et al., 1975, 2006), political views (Lin & Ensel, 1981, Fond & Neville, 2010), preferences of music genre (Rentfrow & Gosling, 2003), and emotional reaction to music genres. This approach provided the opportunity to reinforce the idea on how psychology and social networks studies are intertwined. The analysis of the main results obtained from the three empirical studies is shown below.

a) The results provided information on user interactions that can be used in personalized recommendation. This information is based on structural dimensions related to the users' location in the network, and with their personal attributes. This can be applied to create a representation of the users and their social links in the network in relation to other users, with whom the user (the receiver of recommendation) would have a weak and non-redundant tie while forming a network bridge. Thus, the receiver could be surprised by the recommendation delivered. Therefore, given the theoretical approach discussed in the fourth chapter, I believe that this kind of social data can counteract the Social Echo Chamber Effect. Furthermore, it allows us to speculate about the added value for receivers; first, when they interpret information based on novelty – notably by supporting a richer construction of meaning due to subconscious activity; second, by the gain in affect through the elicited emotion, i.e., surprise. This is important in the scope of recommendation, but can also be applied in other fields of analysis, like searching;

b) The empirical results offer promising evidence about the relevance of the study of emotions in the context of SNAs. Three synthesizing principles guide this overview.

First, *surprise is an adequate proxy of the novelty perceived by a receiver in a network bridge*. The main contribution of Chapter 6 is the proposed method – surprise as proxy of perceived novelty – which identifies the relationship between bridging assumptions and the perception of novelty. However, as the proposed method demonstrates, not all bridges assuming the delivery of novelty can match the receivers' perception of novelty. Burt (1992) proposes a measure to calculate bridges, and so, to find the bridging assumptions related to the delivery of novelty. Nevertheless, this theory does not explain whether there is a perception of novelty or not. A similar constraint can be found in the bridging theory of Granovetter (1973). This method extends the results on theories of bridging by introducing the receiver's viewpoint – their perception of novelty. A valuable contribution to the SNA field was obtained testing the methods. A difference was found between the number of bridges that can be assumed by following the traditional approach and the one found by means of the perceived novelty.

The findings have shown that socially distant ties and a low emotional support between sender and receiver are two important dimensions to describe weak ties as bridges. On the other hand, a bridge spanning a structural hole is considered non-redundant only if this link does not contain more than one strong tie in the triads formed with common third-party connections. Furthermore, no-redundancy is more prevalent in bridges connecting weakly tied individuals. Lastly, network bridges are an important structural condition to explain the emotional reaction of surprise.

A second synthesizing principle is that *structural holes spanned by weak ties reveal a strong relationship with receivers' choices of contents.* Content selection is more dependent on the receiver's emotional reaction (i.e., surprise), than on factors associated with the node's social relevance. Social relevance means here the number of adjacent connections (degree centrality) and corresponding tie strength of a given node relative to other nodes in the network. This argument is the main contribution of the seventh chapter. Chapter 7 also shows that centrality (degree and betweenness) and strength of ties (i.e., friendship) are less relevant for content selection than has been discussed in literature, notably, given the value attributed to popularity and friendship.

Literature usually debates the benefit of individuals in the central position. In this context, when the central position is defined by centrality measures, the degree is the base of assessment. It means that such central positions will be correlated with the benefits associated to the centrality degree, which often includes the grade of popularity, prestige, or influence (Knoke & Burt, 1983). Regarding to betweenness centrality, the corresponding benefit of such a network position is accounted by the ability to broke the flow of information.

On the other hand, when this central position is identified by a bridge, the benefit is centered on the access to novelty, in other words on the ability of the receiver to be surprised – as the sixth chapter shows. However, contrary to the proposal of this dissertation, the receiver's perspective (e.g., content choices) is not usually discussed in literature, nor the benefits (at cognitive level) when the information is received. As presented in the second chapter, there are cognitive gains related to the perception of novelty, which stimulates a richer construction of meaning. This happens because surprise is an emotion stimulated at an unconscious level and so is relevant in promoting the use of implicit information in the interpreting processes. Therefore, the study also contributes to a clarification of the differences between sender and receiver when a content selection is made and, additionally, to characterize the corresponding gains. Finally, it highlights the relevance of network bridges as central nodes that determine the information flow in a network.

Third, the results suggest that *personal attributes (differences in gender) jointly with bridging factors (weak ties) characterize the optimal cognitive distance (between individuals in a social network) underlying the perception of novelty, i.e., surprise.* It means that such dimensions are a predictor of the perception of novelty. This is the chief contribution of the eighth chapter. Although transitivity mechanisms, i.e., based on homophily dynamics that traditionally underlie social mechanisms of triadic closure and selective mixing (Goodreau et al., 2009), have been applied to provide targeted product recommendations, this study shows that is the heterophily in gender that explains surprise, rather than homophily, as might be expected. As a matter of fact, among the range of personal attributes tested, it is dissimilarity in gender (heterophily) that predicts surprise. This happens 74% of the times that surprise is elicited.

Heterophily describes the tendency to interact with others of different type and refers to the fact that different people can have different frequencies or intensities in their relations. E.g., there is homophilia among the members of the same social class and there is heterophily among members of different class or age. Because the study of heterophily has not been approached very often in sociology, this result seems to be important, notably to communication studies.

The approach tested analyzes the association between personal attributes and bridging outputs to identify the best combining between structural conditions and the adequate cognitive distance among users that assure better odds of novelty perception. In this regard, it is assumed that bridges can be seen as enablers that approach people with mutual interest on similar topics – but beyond the effects of echo chamber. On the other hand, the perception of novelty is conceived in a framework of communication because surprise requires a certain level of agreement between sender and receiver. The findings seem to support this idea, but they also highlight that the endogenous effects, like homophily and similarities based on music and political interests, do not fully characterize such communication process.

The findings also contribute to raising the assumption that people shape their networks (not the other way around) depending on perception of others through the access to content and its assessment at cognitive and emotional level. Contents can be seen here as a proxy that interfaces the emotional and affective (virtual) contact between sender and receiver. Therefore, the results seem to support the claim that the cognitive stimulus related to the interpretation of information is not only dependent on the information itself, but also on the emotions elicited by individuals. In particular, there is an optimal cognitive distance between sender and receiver when the surprise is the elicited emotion. As seen, this distance can be characterized by the individuals' structural position in the network and their personal attributes.

In summary, this dissertation characterizes the problem of the Social Echo Chamber Effect, which affects online users when they receive a personalized recommendation online. Because this problem arises due to the use of social data, it was developed a study focused on an alternative extraction of social data. It consisted of three empirical studies following a simple premise: surprise as a proxy of novelty to study the individual attributes and network dimensions that explain the perception of novelty by receivers in a network environment. Thus, by introducing the study of emotions (i.e., surprise response), the flow of information, which is usually weighted by the location and number of network members (Shi et al., 2013) and regarded as structurally (Burt, 2002) and content dependent (Holme & Ghoshal, 2008) is, additionally, emotionally weighted. This presents a new perspective on the role of emotions' in social networks. Therefore, given the definition that a network is a set of relationships (Kadushin, 2004), while a social network "consists of a finite set or sets of actors and the relation of relations defined in them" (Wasserman & Faust, 1994, p.20), it might be correct to add that social networks are also users' choices in response to elicited emotions.

Moreover, the results show that social data can be used in a way that increases the cognitive distances among users in order to deal with a set of new threats that has been imposed by some web algorithms. Some of such threats can be named as: a) *Democracy or Tolerance threats*, because people are being separated by opinion clusters[75]; b) *Conformity threat*, given the lack of "natural" liberty to access novel information; c) *Cognitive threat*, given that people's ability to interpret surrounding reality is diminished; d) *"Fluffy" Innovation threat*, due to the urgency to obtain people's time and attention, which can reduce the added value to society of some technologies.

Despite several results that point towards interesting outcomes for the three fields of study covered by this work, the general conclusion is that *the performance of social network-based recommender systems can be improved through social data conceived from differences in gender and central nodes defined by network bridges of distant ties spanning non-redundant structural holes.* Such dimensions defines the optimal cognitive distance between users (i.e., sender and receiver of recommendation) related to novelty perception. Non-redundancy means having no more than one strong tie between the triads formed by sender, receiver and common third-parties. This underscores the idea that receivers of such recommendations will benefit from the novelty delivered, but also from a richer construction of meaning due to a subconscious cognitive process that is stimulated by novelty interpretation and so, by the emotion of surprise.

[75] Ethan Zuckerman, director of the center for Civic Media at MIT. See: https://www.bbvaopenmind.com/ethan-zuckerman-todavia-no-entendemos-muy-bien-como-ocurre-el-cambio-social-en-la-era-digital/.

In closing, it is important to acknowledge some of the shortcomings of this project. In this regard, it should be noted that the dimension of the sample used may weaken a more general view about the results obtained. Despite the multiple assessments undertaken, several ties were outside the observation range, which may have hidden some information on third party-connections forming triadic relationships among participants not detected through the data gathered. Thus, it would be relevant, first, to test the methods proposed in this dissertation with a larger population; second, to develop this experiment with a population from a Recommender System; third, to find further and stronger evidence of regularities in the associations tested between personal attributes, network bridges and surprise would strengthen the findings of the proposed method.

Furthermore, the interdisciplinary approach of this study enables a contribution to three scientific fields: 1) Social Network Analysis; 2) Recommender Systems; and 3) Social-psychology.

1) The contribution for the Social Network Analysis field is mainly focused on the method proposed for analyzing bridging assumptions and relationships between network dimensions and emotional response, particularly surprise. This is relevant in the sense that it contributes from a different perspective to the study of weak ties and structural holes. On the other hand, this study draws attention to the importance of developing more interdisciplinary work between both fields of social-psychology and SNA.

2) The contribution to Recommender Systems is twofold. First, we framed the problem related to personalization in the context of Social Echo Chamber Effect, and explained how the potential of innovation in Web technologies can be compromised by cognitive factors. Second, I discussed a solution for this problem based on the use of specific data from users' social networks. This dissertation ends with a theoretical proposal applied to social-based recommender system using the empirical results of this investigation. It draws our attention to the possibility of delivering novel and surprising recommendations based on prediction, instead of randomly. The next steps would be to apply the findings in the development of an algorithm and to test it on a social network - based recommender system.

3) The last contribution of this dissertation is aimed at Social-psychological studies. In this scope, this work deliberates on how to enrich the construction of

meaning of the target-users of novel information by means of network dimensions and personal attributes. It also shows the importance of having further contributions from this field of studies to develop the understanding of social networks from the viewpoint of their actors, rather than be seen eminently from a structural perspective. Despite some important contributions from Social-psychological studies in the field of Social Network Analysis, further studies applied to Digital Media Systems are needed. It would be particularly interesting to find new associations between personal attributes and surprise response namely in the context of bridging factors (i.e., weak ties and non-redundant structural holes). Additionally, it would be important to test personal attributes that could be extracted directly, or implicitly, from network social data.

LIST OF REFERENCES

Abbasi, A., & Altmann, J. (2010). A Social Network System for Analyzing Publication Activities of Researchers. Symposium on Collective Intelligence, COLLIN 2010, in: *Advances in Intelligent and Soft Computing*, Springer, Hagen, Germany.

Aczel, J., & Daroczy, Z. (1975). *On measures of information and their characterizations*. Academic Press, New York.

Adomavicius, G., & Tuzhilin, A. (2011). Context-Aware Recommender Systems. Ricci, F., Rokach, L., Shapira, B., & Kantor, P., B. (eds.) *Recommnder System Handbook*. Springer Science+Business Media, LLC 2011.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, Nr. 6.

Agichtein, E., Brill, E., Dumais, S. & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. In Proceedings of *SIGIR*, pages 3–10. *ACM* Press.

Aggarwal, I. (2013). Cognitive Style Diversity in Teams. Business Commons. Dissertations, Paper 258, from http://repository.cmu.edu/dissertations.

Ahuja, G., Soda, G., & Zaheer, K. (2012). The Genesis and Dynamics of Organizational Networks. *Organization Science*, Vol. 23 Issue 2, March-April, pp. 434-448.

Ali, K., & van Stam, W. (2004). TiVo: making show recommendations using a distributed collaborative filtering architecture. Proceedings *10th ACM SIGKDD Int Conf Knowl Disc Data Min* 394–401.

Anagnostopoulos, A., Kumar, R., & Mahdian, M. (2008). Influence and Correlation in Social Networks. *KDD'08*, August 24–27, Las Vegas, Nevada, USA, ACM 978-1-60558-193-4/08/08.

Anderson, C. (2006). *The Long Tail: Why the Future of Business is Selling Less of More*, Hyperion.

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proceedings of the *National Academy Science U S A*. 2009 December 22; 106(51): 21544–21549.

Aral, S., & Alstyne, M. (2011). The Diversity-Bandwidth Trade-off. *American Journal of Sociology*, Vol. 117, No. 1 (July 2011), pp. 90-171.

Arazy, O., Kumar, N., & Shapira, B. (2010). A Theory-Driven Design Framework for Social Recommender Systems. *Journal of the Association for Information Systems* 11(9), pp. 455–490.

Armstrong, S. (1999). The influence of individual cognitive style on performance in management education. In J. Hill, S. Armstrong, M. Graff, S. Rayner, & E. Sadler-Smith, (Eds.), Proceedings of the *4th Annual Conference of the European Learning Styles*.

Armstrong, S.J., Cools, E., & Sadler-Smith, E. (2012). Role of Cognitive Styles in Business and Management: Reviewing 40 Years of Research. *International Journal of Management Reviews*, Vol. 14, 238–262.

Arora G., Kumar, A., Devre, G.S., & Ghumare, A. (2014). Movie recommendation system based on users' similarity. *International Journal of Computer Science and Mobile Computing*, Vol. 3, Issue. 4, April, p.765 – 770.

Back, M.D., Stopfer, J.M., Vazire, S., Gaddis, S., Schmukle, S.C., Egloff, B., & Gosling, S.D. (2010). Facebook Profiles Reflect Actual Personality, Not Self-Idealization. *Psychological Science* 21(3) 372–374.

Backstrom, L., Huttenlocher, D., Kleinberg, J., & Lan. X. (2006). Group formation in large social networks: membership, growth, and evolution. In KDD '06: Proceedings of the 12th ACM SIGKDD, *International conference on Knowledge discovery and data mining*.

Bakshy, E., Rosenn, I., Marlow, C., & Adamic., L. (2012). The Role of Social Networks in Information Diffusion. *WWW 2012*, April 16–20, Lyon, France, ACM 978-1-4503-1229-5/12/04.

Baldi, P., & Itti; L. (2010). Of Bits and Wows: A Bayesian Theory of Surprise with Applications to Attention. *Neural Netw*ork, 2010 June, 23(5): 649–666.

Bargh, J.A., & Morsella, E. (2008). The Unconscious Mind. *Perspective Psychology Science*. 2008 January ; 3(1): 73–79.

Bargh, J. (2011). Unconscious Thought Theory and Its Discontents: A Critique of the Critiques. *Social Cognition*, Vol. 29, No. 6, 2011, pp. 629–647.

Barsade, S. G., Brief, A. P., & Spataro, S. E. (2003). The affective revolution in organizational behavior: The emergence of a new paradigm. In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (2nd ed.). Mahwah, NJ: Erlbaum.

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or Surprise? *Frontiers in psychology*, Vol.4, 907-1.

Bateson, G. (2000). *Steps to an Ecology of Mind*. University of Chicago Press, Chicago [Originally published in 1972].

Batorski, D. (2012). The Burden of Big Data. New technologies for generating giant databases with applications in business and science. From: http://www.academia.edu/4327496/The_Burden_of_Big_Data._New_technologies_for_generating_giant_databases_with_applications_in_business_and_science.

Bauerfeind, U. (2003). The evaluation of a recommendation system for tourist destination decision making. In Proceedings of the *XII International Symposium on Tourism and Leisure*, Barcelona, 3-4 April.

Bender, M., Crecelius, T., Kacimi, M., Michel, S., Neumann, T., Parreira, J.X., Schenkel, R., & Weikum, G. (2008). Exploiting social relations for query expansion and result ranking. In Proceedings of ICDE Workshops, pages 501–506. *IEEE*.

Berger, P.L., Luckmann, T. (1990). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor Books, New York [1966].

Berners-Lee, T., Fischetti, M., Francisco, H. (1999). Weaving the Web: the Original Design and Ultimate Destiny of the World Wide Web by its Inventor. *Harper*, San Francisco, CA.

Berlyne, D.E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw Hill.

Berlyne, D.E. (1965). *Structure and direction in thinking*. New York: Wiley.

Bilgic, M. (2004). *Explanation for Recommender Systems: Satisfaction vs. Promotion*. Undergraduate Honor Thesis, May 2004.

Blau, J. (1980). *When Weak Ties Are Structured*. Unpublished manuscript, Department of Sociology, State University of New York, Albany.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. Elsevier, *Knowledge-Based Systems*, Vol. 46, July, Pages 109–132.

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology* 92:1170-1182.

Bonacich, B. (1991). Simultaneous group and individual centralities. *Social Networks*, Vol. 13, Issue 2, June 1991, Pages 155–168.

Bohn, A., Buchta, C., Hornik, K., & Mair, P. (2014). Making friends and communicating on Facebook: Implications for the access to social capital. *Social Networks*, 37, 29–41.

Bonhard, P. (2004). Improving Recommender Systems with Social Networking. CSCW04, from http://www.cs.ucl.ac.uk/staff/p.bonhard/pubs/CSCW04.pdf .

Bonhard, P., Sasse, & M.A. (2006). Knowing me, knowing you' – using profiles and social networking to improve recommender systems. *BT Technology Journal*, Vol.24, No 3.

Borgatti, S.P., (2005). Centrality and network flow. *Social Networks*, 27 (2005) 55–71.

Borgatti, S.P., & Foster, P.C. (2003). The Network Paradigm in Organizational Research: A Review and Typology. *Journal of Management*, 29(6) 991–1013.

Borgatti, S.P., Everett, M.G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, 28 (2006) 466–484.

Borgatti, S. P., Carley, K., & Krackhardt, D. (2006). Robustness of centrality measures under conditions of imperfect data. *Social Networks,* 28 (2), 124-136.

Borgatti, S.P., & Everett, M.G. (2009). Models of corer/ periphery structures. *Social Networks*, 21 1999. 375–395.

Borgatti, S.P., & Lopez-Kidwell, V. (2011). in The SAGE HandBook: SNA. (Eds) *Network Theory*. Sage Publications, London.

Bossche, P.T., Gijselaers, W.H., Segers, M., & Kirschner, P.A. (2006). Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments: Team Learning Beliefs and Behaviors. *Sage*, Vol. 37, Nr. 5, October, 490-521.

Boyd, D. (2002). *Faceted id/entity: Managing representation in a digital world*. Master's thesis, MIT.

Boyd, D.M., & Ellison, N.B. (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13, 210-230, International Communication Association, from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2007.00393.x/full.

Brass, D.J. (1984). Being in the right place: A structural analysis of individual influence in an organization. *Administrative Science*, Quarterly 29, 518-539.

Brass, D. (1985). Men's and women's networks: A study of interaction pattern's and influence in an organization. *Academy of Management Journal*, 28, 327-343.

Brass, D. J., & Burkhardt, M.E. (1992). Centrality and power in organizations. In N. Nohria & R.G. Eccles, (Eds.), *Networks and organizations: Structure, form, and action* (pp. 191-215). Boston: Harvard Business School Press.

Brass, D.J. (2011). A social network perspective on industrial/organizational psychology. In S. Koslowski, (Ed.), *Handbook of industrial and organizational psychology,* Vol.1.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. In Proceedings of the *Seventh International Conference on World Wide Web*, volume, pages 107-117.

Breslin, J.G. (2009). *Semantic Digital Libraries*. Part 1 – Social Semantic Information Spaces. Springer. Ireland.

Bruner, J. (1986). *Actual Minds, Possible Worlds*. Harvard University Press, Cambridge, MA.

Brynjolfsson, J., Hu, Y., Smith, & M.D. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science* 49 (2003) 1580-1596.

Bulte, V.D., & C., Wuyts, S. (2007). *Social networks and marketing*. Marketing Science Institute, Cambridge, MA.

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. User Model. *User Modeling and User-Adapted Interaction*, 12 (2002) 331-370.

Burt, R. S. (1992). *Structural Holes*. Harvard University Press: Boston.

Burt, R. (1995). *Structural Holes: The Social Structure of Competition*. Harvard University Press.

Burt, R. S., Jannotta, J. E., & Mahoney, J. T. (1998). Personality correlates of structural holes. *Social Networks*, 20, 63–87.

Burt, R. S. (2002). The social capital of structural holes. In M. F. Guillen, R. Collins, P. England, & M. Meyer, (Eds.), *The new economic sociology* (pp. 148 –189). New York: Russell Sage Foundation.

Burt, R.S. (2002). The Social Capital of Structural Holes. In Mauro F. Guillén, Randall Collins, Paula England, and Marshall Meyer, (Eds), *New Directions in Economic Sociology*, New York: Russell Sage Foundation.

Burt, R.S. (2004). Structural Holes and Good Ideas. *AJS*, Vol. 110, Nr. 2, September, 349–99.

Burt, R.S. (2007). Second-hand brokerage: Evidence on the importance of local structure for managers, bankers, and analysts. *Academy of Management Journal*, 50, 110-145.

Butts, C., & Carley, K. (1999). Spatial Models of Network Formation. *Working Paper*, 1–36.

Cai, Y., Yoon, Y., & Kim, W. (2014). Personalized Search System Based on User Profile. *Springer*, pp 320-328.

Carmel, D., Zwerdling, N., Guy, I., Ofek-Koifman, S., Har'el, N., Ronen, I., Uziel, E., Yogev, S., & Chernov, S. (2009). Personalized social search based on the user's social network. *CIKM '09*, Proceedings of the *18th ACM conference on Information and knowledge management*, pages 1227-1236.

Casati, R., & Pasquinelli, E. (2007). How can you be surprised? The case for volatile expectations. *Phenomenology and the Cognitive Sciences*, 6 (1-2), 171-183(13).

Casciaro, T., Lobo, M.S. (2008). When competence is irrelevant: the role of interpersonal affect in task-related ties. *Administrative Science Quartly*, 53:655–84.

Centola, D. (2007). Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups. Sage Publications, *Journal of Conflict Resolution*, Volume 51, Number 6, 905-929.

Centola, D. (2010). The Spread of Behavior in an Online Social Network Experiment. Science, 329(5996): 1194-1197, September.

Christakis, N.A., & Fowler, J.H. (2009) *Connected: The Surprise Power of Our Network and How they Shape Our Lives*. 1st Edition, Little, Brown and Company, New York.

Chandler, D. (2005). Semiotics for Beginners: Codes. Retrieved 11 March 2011, from http://www.aber.ac.uk/media/Documents/short/whorf.html.

Chang, C-C. (2013). Examining users0 intention to continue using social network games: A flow experience perspective. *Telematics and Informatics*, 30, 311–321.

Chen, L. & Sycara, K. (1998). WebMate: A Personal Agent for Browsing and Searching. In Proceedings of the *Second International Conference on Autonomous agents*, ACM, AGENTS '98.

Chen, H. (2000). Exploring Web users' optimal flow experiences. *Information Technology & People*, Vol. 13 No. 4, 2000, pp. 263-281.

Chen, C.M., & Sun, Y.C. (2012). Assessing the effects of different multimedia materials on emotions and learning performance for visual and verbal style learners. *Computers & Education*, 59(4), 1273–1285. http://www.sciencedirect.com/science/article/pii/S187704280900158X

Chen, S.Y. and Macredie, R.D. (2002). Cognitive modelling of student learning in web-based instructional programs. *International Journal of Human–Computer Interaction*, 17, pp. 375–402.

Chiu, C.M., Hsu, M.H., & Wang, E.T.G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, 42(3), 1872–1888.

Cohen, J. (2001). Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication & Society*, 4(3), 245–264.

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *American Journal of Sociology*, 94:S95–S120.

Collier, J. (1990). Intrinsic information, in: P.P. Hanson (Ed.), *Information, Language, and Cognition*: vol. 1, *Vancouver Studies in Cognitive Science* (pp. 390–409). Oxford UP.

Constant, D., Sproull, L., & Kiesler, S. (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science*, 7(2), 119–135.

Cook, D.A. (2005). Learning and cognitive styles in web-based learning: theory, evidence and application. *Academic Medicine*, 83, pp. 266–278.

Cornwell, B. (2009). Good health and the bridging of structural holes. *Social Networks*, 31 (2009) 92–103.

Cover, T.M., & Thomas, J.A. (1991). *Elements of Information Theory*. John Wiley, New York.

Crosier, B.S., Webster, G.D., & Dillon, H.M. (2012). Wired to Connect: Evolutionary Psychology and Social Networks. *Review of General Psychology*, Vol. 16, No. 2, 230–239.

Davidson, D. (2004). *Problems of rationality*. Oxford: Oxford University Press.

Darwin, C. (1859). *On the origin of species*. John Murray; London.

Davidson, D. (2004). *Problems of rationality*. Oxford: Oxford University Press.

Dennett, D.C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. Simon & Schuster; New York.

Derbaix, C., & Vanhamme, J. (2003). Inducing word-of-mouth by eliciting surprise – a pilot investigation. *Journal of Economic Psychology*, vol. 24, Issue: 1, pp. 99-116.

D'Esposito, M., & Zaccarin, S. (2011). Editorial: applied and methodological issues in the analysis of network data. *Quality & Quantity* 45, 985–987 (2011),.

Dewey, J. (1876). *A Classification and subject index cataloguing and arranging the books and pamphlets of a library*. Amherst, Mass. Facsimile reprinted by Forest Press Division Lake Placid Educational Foundation. Retrieved 22 March 2011, from http://www.gutenberg.org/files/12513/12513-h/12513-h.htm.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, 1, 95–109.

Ding, Y. (2001). A review of ontologies with the Semantic Web in view. *Journal of Information Science*, vol. 27 (6), pp. 377–384.

Ding, Y., 2010. Semantic Web: Who is in the field – a bibliometric analysies. *Journal of Information Science*, vol. 36 (3), pp. 335–356.

Dou, Z., Song, R., & Wen, J.R. (2007). A large-scale evaluation and analysis of personalized search strategies. In Proceedings of *WWW*, pages 581–590. *ACM*.

Driver, B.L., Brousseau, K.R., & Hunsaker, P.L. (1990). The dynamic decision-maker: Five decision styles for executive and business success, *Harper & Row*, New York.

Duncan, G.J., & Raudenbush, S.W. (1998). Neighborhoods and Adolescent Development: How Can We Determine the Links?. in Does It Take a Village? Community Effects on Children, Adolescents, and Families. A. Booth and A. C. Crouter, (Eds.), *Pennsylvania State University Press* (pp. 105–136). State College, Pa, USA, from http://www.gse.uci.edu/person/duncan_g/docs/neighborhoods.pdf.

Eco, U. (1990). O*s limites da interpretação*. Carnaxide, Difel.

Ekman, P., & Friesen, W.V. (1975). *Unmasking the face. A guide to recognizing emotions from facial clues*. Englewood Cliffs, N.J., Prentice-Hall Inc.

Ekman, P., & Davidson, R.J. (Eds.) (1994). *The Nature of Emotion: Fundamental Questions*. Oxford: Oxford University Press.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, 3, 364–370.

Ekstrom, S.R. (2004). The mind beyond our immediate awareness: Freudian, Jungian, and cognitive models of the unconscious. *Journal of Analytical Psychology,* 2004, 49, 657–682.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook ‘‘friend’’: Social capital and college students’ use of online social network sites. *Journal of Computer – Mediated Communication*, 12(4), 1143–1168.

Entman, R.M. (1989). How the media affect what people think: An information processing approach. *The journal of Politics*, Vol. 51, Issue 2, May, 347-370.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, 49, 709–724.

Epstein, S. (2003). *Cognitive-experiential self-theory of personality*. In T. Millon, M. J. Lerner, & I. B.Weiner, (Eds.), *Comprehensive handbook of psychology: Personality and social psychology*, Vol. 5 (pp. 159–184). Hoboken, NJ: Wiley.

Evans, V. (2006). Lexical Concepts, Cognitive Models and Meaning-Construction. *Cognitive Linguistics*, 17(4), 491–534.

Fecteau, J.H., Kingstone, A., & Enns, J.T. (2004). Hemisphere differences in conscious and unconscious word reading. Elsevier, *Consciousness and Cognition*, 13, 550–564.

Feld, S.L. (1981). The Focused Organization of Social Ties. *The American Journal of Sociology*, vol. 86, No. 5, pp. 1015-1035.

Fernandez-Luque, L., Karlsen, R. & Vognild, L.K. (2009). Challenges and Opportunities of Using Recommender Systems for Personalized Health Education. *In proceeding of: Medical Informatics in a United and Healthy Europe - Proceedings of MIE 2009, The XXIInd International Congress of the European Federation for Medical Informatics*, Agust 30 - September 2, Sarajevo, Bosnia and Herzegovina.

Fernández, Y.B., Arias, J.J. P., Solla, A.G., Cabrer, M.R., & Nores, M.L. (2006). Bringing together content-based methods, Collaborative filtering and semantic inference to improve personalized TV. *4th European Conference on Interactive Television* (EuroITV 2006).

Ferris, G.R., & Treadway, D.C. (2012). *Politics in Organizations. Theory and Research Considerations*. Routledge & Francis Group. New York.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Ficher, M. (2010). *Birds of a feather flock together reloaded: Homophily in the Context of Web 2.0 in Online Networking Sites such as Facebook.* Master Thesis. The Graduate School of The College of Charleston, USA.

Fillmore, C. J. (1976). The need for frame semantics within linguistics. *Statistical Methods in Linguistics*, 12(5), 5–29.

Finkenauer, C., Luminet, O., Gisle, L., El-Ahamadi, A., Linden, V. D. M., & Philippot, P. (1998). Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model. *Memory & Cognition*, 26 (3), 516-531.

Finneran, C.M., & Zhang, P. (2005). Flow in Computer-Mediated Environments: Promises and Challenges. *Communications of the Association for Information Systems*, Vol. 15, Article 4.

Flavián-Blanco, C., Gurrea-Sarasa, R., & Orús-Sanclemente, C. (2013). Analyzing the emotional outcomes of the online search behavior with search engines. *Computers in Human Behavior*, 27, 540–551.

Fleming, L., Mingo, S., Chen, D. (2007). Collaborative Brokerage, Generative Creativity and Creative Success. *Administrative Science Quarterly*, 52: 443–475.

Freeman, L. C. (1979). Centrality in networks: I. conceptual clarification. *Social Networks*, 1, 215–239.

Fond. T., Neville, J. (2010). Randomization Tests for Distinguishing Social Influence and Homophily Effects. *WWW 2010*, April 26–30, Raleigh, North Carolina, USA.

Fowler, J.H. & Christakis, N.A. 2008. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, 337, no. a2338: 1-9

Fowler, J.H., Dawes, C.T.,& Christakis, N.A. (2009). Model of genetic variation in human social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 1720–1724.

Fredrickson, B. L. (1998). What good are positive emotions?. *Review of General Psychology*, 2, 300-319.

Fredrickson, B. L. (2000). Cultivating positive emotions to optimize health and well-being. Target article in Prevention and Treatment, 3, from http://journals.apa.org/prevention.

Freeman, W. (2003). A Neurobiological, theory of Meaning in Perception. Part I: Information and meaning in Nonconvergent and nonlocal Brain Dynamics. *International Journal of Bifurcation and Chaos*, Vol. 13, No. 9 (2003) 2493–2511.

Friedkin, N. (1980). A Test of Structural features of Granovetter's Strength of weak Ties Theory. *Social Networks*, 2 (1980) 411-422.

Frijda, N.H. (2003) Emotion experience. *Cognition and Emotion*, 0(0), CEM 1450, from http://linus.media.unisi.it/cirg/contact/frijda05.pdf.

Furnham, A. (2011). *Intelligence and intellectual Styles*. (Eds) Zhang, L.-F., Sternberg, R.J. & Rayner, S. (2012) Handbook of Intellectual Styles: Preferences in Cognition, Learning, and Thinking. Springer Publishing Company.

Gainotti, G. (2012). Unconscious processing of emotions and the right hemisphere. *Neuropsychologia*. Volume 50, 205–18.

Gasparetti, F., Micarelli, A., & Sansonetti, G. (2014). Mining Navigation Histories for User Need Recognition. *Springer*, Vol. 434, pp 169-173.

Gilbert, E., & Karahalios, K. (2009). Predicting Tie Strength With Social Media. *CHI 2009*, April 4–9. Boston, Massachusetts, USA.

Gilbert, E. (2012). Predicting Tie Strength in a New Medium. *CSCW'12*, February 11–15. Seattle, Washington, USA. ACM 978-1-4503-1086-4/12/02.

Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., & Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. Elsevier, *Research Policy* 37, 1717–1731.

Golbeck, J., & Wasser, M.M. (2007). Social Browsing: Integrating Social Networks and Web Browsing. *ACM*, 9 78-1-59593-642-4/07/0004.

Golder, S.A. & Yardi, S. (2010). Structural Predictors of Tie Formation in Twitter: Transitivity and Mutuality. *Proceedings of the Second IEEE International Conference on Social Computing*, August 20-22. Minneapolis, MN.

Goldsmith, M. (1934). *A history of mesmerism*. Doubleday, Doran & Company; Garden City, New York.

Goodreau, S.M., Kitts, J.A., & Morris, M. (2009). Birds of a Feather, or friend a friend? Using exponential random models to investigate adolescent social networks. *Demography*, vol. 46, Nº 1, February 2009: 103–125.

Guo, G., Zhang, J., Thalmann, D., Basu, A., & Yorke-Smith, N. (2014). From Ratings to Trust: an Empirical Study of Implicit Trust in Recommender Systems. *SAC'14*, March 24-28, Gyeongju, Korea.

Gosling, S.D., Vazire, S., Srvastava, S., & John, O.P. (2004). Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions about Internet Questionnaires. *American Psychological Association*. Vol. 59, No. 2, 93–104.

Graells-Garrido, E., Lalmas, M., & Quercia, D. (2013). Data Portraits: Connecting People of Opposing Views. from: http://arxiv.org/pdf/1311.4658v1.pdf.

Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*, vol. 78, Nº 6, pp. 1360 – 1380.

Granovetter. M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociology Theory*, Vol. 1, pp. 201-233.

Granovetter. M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *The American Journal of Sociology*, Vol. 91, No. 3 (Nov., 1985), pp. 481-510.

Gretzel, U., Hwang, Y-H., & Fesenmaier, D. (2012). Informing destination recommender systems design and evaluation through quantitative research. International Journal of Culture, *Tourism and Hospitality Research*, 6 (4), 297-315.

Groh, G., Ehmig, C. (2007). Recommendations in Taste Related Domains: Collaborative Filtering vs. Social Filtering. *GROUP'07*, November 4–7, Sanibel Island, Florida, USA.

Gruber, T.R. (1995). Towards principles for the design of ontologies used for knowledge sharing. *International Journal of Human–Computer Studies*, 43 (5/6) 907–28.

Gu, B., Konana, P., & Chen, M. (2008). Melting-Pot or Homophily? – An Empirical Investigation of User Interactions in Virtual Investment-Related Communities. *Working Paper*, University of Texas.

Guo, G., Zhang, J., Thalmann, D., Basu, A., & Yorke-Smith, N. (2014). From Ratings to Trust: an Empirical Study of Implicit Trust in Recommender Systems. *SAC'14* March 24-28, Gyeongju, Korea.

Haas, S.A., Schaefer, D.R., & Kornienko, O. (2010). Health and the Structure of Adolescent Social Networks. *Journal of Health and Social Behavior,* 51: 424.

Hampton, K.N., Sessionsa, L.F., & Her, E.J. (2010). Core Networks, Social Isolation, and New Media – How Internet and mobile phone use is related to network size and diversity. *Communication & Society,* 14: 1, 130 – 155.

Han, L., & Yan, L.H. (2009). A fuzzy biclustering algorithm for social annotations. *Journal of Information Science*, Vol. 35 no. 4 426-438

Hansen, M.T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44, 82-111.

Hansen, M. T. (2002). Knowledge Networks: Explaining Effective Knowledge Sharing in Multiunit Companies. *Organization Science*, 13 (3): 232–48.

Harrigan, N., Achananuparp, P., & Lim, E. (2012). Influentials, novelty, and social contagion. The viral power of average friends, close communities, and old news. *Social Networks,* 34 (2012) 470– 480.

Haythornthwaite, C. (1996). Social Network Analysis: An Approach and Technique for the Study of Information Exchange. *Library & Information Science Research*, Vol.18, Issue 4, Autumn 1996, Pages 323–342.

He, J., Chu, W. W. (2010). A Social Network-Based Recommender System (SNRS). *Annals of Information Systems,* 12 (2010) 47-74.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

Heine, S.J., Proulx, T., & Vohs, K.D. (2006). The Meaning Maintenance Model: On the Coherence of Social Motivations. *Personality and Social Psychology Review*, Vol.10, No.2 , 88-110.

Heinström, J. (2006). Psychological factors behind incidental information acquisition. *Library & Information Science Research*, Vol. 28, Issue 4, Winter 2006, Pages 579-594.

Hipp, J.R., & Perrin, A.J. (2009). The Simultaneous Effect of Social Distance and Physical Distance on the Formation of Neighborhood Ties. *City & Community*, 8:5-25.

Hoffman, D.L., Novak, T.P. (2009). Flow Online: Lessons Learned and Future Prospects. *Journal of Interactive Marketing*, 23, 23–34.

Hogg, T., Wilkinson, D.M., Szabo, G., & Brzozowski, M.J. (2008). Multiple Relationship Types in Online Communities and Social Networks. In Proc. *AAAI Spring Symposium on Social Information Processing*.

Hogarth, R. (2002). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. *Universitat Pompeu Fabra conomics and Business,* Working Paper No. 654.

Holme, P., & Ghoshal, G. (2008). The diplomat's dilemma: Maximal power for minimal effort in social networks. Retrieved July 2008, from http://arxiv.org/abs/0805.3909.

Holtgraves, M., Kashima, Y. (2008). Language, Meaning, and Social Cognition. Sage, *Society for Personality and Social Psychology*, Vol.12: 73.

Homans, G.C. (1958). Social behavior as exchange. *American Journal of Sociology*, 63, 597-606.

Horvitz, E. (2007). Reflections on Challenges and Promises of Mixed-Initiative Interaction. *AAAI Magazine* 28, Special Issue on Mixed-Initiative Assistants.

Huang, A. W., & Chuang, T. (2009). Social tagging, online communication, and Peircean semiotics: a conceptual framework. *Journal of Information Science* 35(3):340–357.

Hunt, R.G., Krzystofiak, F.J., Meindl, J.R., & Yousry, A.M. (1989). Cognitive style and decision making. *Organizational Behavior and Human Decision Processes*, Vol.44, pp. 436-453.

Hwang, W, Cho, Y., Zhang, A., & Ramanathan, M. (2008). Bridging Centrality: Identifying Bridging Nodes in Scale-free Networks. Proceedings of the *14th ACM SIGKDD international conference on Knowledge discovery and data mining*, Pages 336-344.

Ingram P, Morris MW. 2007. Do people mix at mixers? Structure, homophily, and the "life of the party." *Administration Science Quarterly*, 52:558–85.

Itti, L., & Baldi, P.F. (2005). A principled approach to detecting surprising events in video. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 631–637, San Diego, CA.

Itti, L., & Baldi, P.F. (2006). Bayesian surprise attracts human attention. In Y. Weiss, B. Schölkopf, and J. Platt, (Eds.), *Advances in Neural Information Processing Systems 18* (NIPS∗2005) (pp. 547–554), Cambridge, MA: MIT Press.

Itti ,L., & Baldi, P.F. (2009). Bayesian surprise attracts human attention. *Vis. Res.* 49, 1295–306. doi:10.1016/j.visres.2008.09.007.

Izard, C.E., Dougherty, F.E., Bloxom, B.M., & Kotsch, W.E. (1974). The differential emotions scale: A method of surveying the subjective experience of discrete emotions. *Unpublished manuscript*, Vanderbilt University.

Izard, C.E. (1977). *Human Emotions.* New York: Plenum Press.

Izard, C.E. (1991). *The Psychology of Emotions*. New York: Plenum Press.

Izard, C.E. (2011). Forms and functions of emotions: Matters of emotion–cognition interactions. *Emotion Review*, 3, 371–378.

Jackson. M.A. (2008). *Social and Economic Networks*. Forthcoming: Princeton University Press.

Jackson, M. O. (2010). An overview of social networks and economic applications. In J. Benhabib, A. Bisin, & M. O. Jackson, (Eds.), *Hand book of Social Economics* (pp. 11–605). Princeton, NJ: Princeton University Press.

Jamieson, K.H., & Cappella, J.N., (2008). *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. New York, NY: Oxford University Press.

Jeong, W-h, Kim, J., Park, D-s, & Kwak, J. (2013). Performance Improvement of a Movie Recommendation System based on Personal Propensity and Secure Collaborative Filtering. *J Inf Process Syst*, Vol.9, No.1, March.

John, O.P., Naumann, L.P., & Soto, C.J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O.P. John, R.W. Robins, & L.A. Pervin, (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York: Guilford.

Jones, R.A. (2003). Mixed Metaphors and Narrative Shifts: Archetypes. *Sage*, Theory Psychology, vol. 13: 651.

Jung, C.G. (1959). *The concept of the collective unconscious*. In H. Read, M. Fordham, & G. Adler (Eds.), The collected works of C.G. Jung (Vol. 9i, 42–53). London: Routledge. (Original work published 1954).

Kadushin, C. (2012). Understanding social networks: Theories, concepts, and findings. *Oxford University Press*, New York.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720.

Kalish, Y., & Robins, G. (2006). Psychological predispositions and network structure: the relationships between individual predispositions, structural holes and network closure. *Social Networks*, 28, 56–84.

Kanovy, I., & Yaari, O. (2011). Model of opinion spreading in social networks. From http://arxiv.org/ftp/arxiv/papers/1106/1106.0872.pdf.

Kao, G., Lei, P-L., & Sun, C-T (2008). Thinking style impacts on Web search strategies. *Computers in Human Behavior*, 24, 1330–1341.

Katz, R., & Tushman, M.L. (1981). An investigation into the managerial roles and career paths of gatekeepers and project supervisors in a major R&D facility. *R&D Management*, 11: 103-110.

Kaya, H. & Alpaslan, F.N. (2010). Using Social Networks to Solve Data Sparsity Problem in One-Class Collaborative Filtering. *Seventh International Conference on Information Technology,* IEEE 978-0-7695-3984-3/10.

Kecskes, I. (2006). On my mind: thoughts about salience, context and figurative language from a second language perspective. *Second Language Research* 22, 2; pp. 1–19, Edward Arnold (Publishers) Ltd.

Keltner, D., & Gross, J.J. (1999). Functional Accounts of Emotions. *Cognition and Emotion*, 13 (5), 467-480.

Kilduff, M., & Krackhardt, D. (1994). Bringing the individual back in: A structural analysis of the internal market for reputation in organizations. *Academy of Management Journal*, 37: 87-108.

Kim, H. & Decker, S. & Breslin, J. (2009). Representing and sharing folksonomies with semantics. *Journal of Information Science Online First*, published on September 24.

Kim, J., Yoo, J., Lim, H., Qiu, H., Kozareva, Z., & Galstyan, A. (2013). Sentiment Prediction Using Collaborative Filtering. In *ICWSM'13*.

King, L.A., & Hicks, J.A. (2006). Narrating the self in the past and the future: Implications for maturity. *Journal of Research in Human Development*, 3, 121–138.

King, L.a., & Hick, J.A. (2009). Detecting and constructing meaning in life events. *The Journal of Positive Psychology*, Vol. 4, No. 5, September, 317–330.

Knobloch-Westerwick, S., Hastall, M.R. (2006). Social Comparisons With News Personae: Selective Exposure to News. Portrayals of Same-Sex and Same-Age Characters. *Communication Research*, Vol. 33, Nr. 4, August, 262-284.

Knoke, D., & Burt, R.S. (1983). Applied network analysis: A methodological introduction. Prominence. In R. S. Burt & M. J. Miner, (Eds.), 195-222. Beverly Hills, CA: Sage.

Koenigstein, N., Koren, Y. (2013). Towards Scalable and Accurate Item-Oriented Recommendations. *RecSys'13*, October 12–16, Hong Kong, China.

Kossinets, G., & Watts, D.J. (2009). Origins of Homophily in an Evolving Social Network. *American Journal of Sociology,* Vol.115, No. 2, pp. 405-450.

Krackhardt, D. (1990). Assessing the Political Landscape: Structure, Cognition, and Power in Organizations. *Administrative Science Quarterly*, Vol.35, No.2. (Jun., 1990), pp. 342-369.

Krackhardt, D. (1992). The strength of strong ties: The importance of Phil os. In N. Kohria & R. Eccles, (Eds.), *Networks and organizations: Structure, form, and action* (pp. 216-239). Boston: Harvard Business School Press.

Kratzer, J., & Lettl, L. (2008). A Social Network Perspective of Lead Users and Creativity: An Empirical Study among Children. *Creativity and Innovation Management*, Vol.17, Nr.1.

Lai, L.S.L., & Turban, E. (2008). Groups formation and operations in the web 2.0 environment and social networks. *Group Decision and Negotiation*, 17, 387–402.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lampe, C., Ellison, N., & Steinfield, C. (2006). A Face(book) in the crowd: Social searching vs. social browsing. Proceedings of *CSCW-2006* (pp. 167–170). New York: ACM Press.

Lassiter, G. D., Lindberg, M. J., Gonzalez-Vallejo, C., Bellezza, F. S., & Phillips, N. D. (2009). The deliberation-without attention effect: Evidence for an artifactual interpretation. *Psychological Science*, 20, 671–675.

Lazarsfeld P.F., & Merton R.K. (1954). Friendship as a social process: a substantive and methodological analysis. In *Freedom and Control in Modern Society*, ed. M. Berger, pp. 18–66.

Lazarus, R.S. (1984). On the primacy of affect. *American Psychologist*, Vol.39 (2), 117-123.

Lazer, D., Pentland , A.S., Adamic, L., Aral, S., Barabasi, A.L, Brewer, D. , Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Alstyne, M.

(2009). Life in the network: the coming age of computational social science. Science. Feb 6, 323(5915): 721–723.

Lee, J., & Kim, S. (2011). Exploring the role os social organizational commitment: Network centrality, strength of ties, and structural holes. *The American Review of Public Administration*, 41(2)205-223.

Leenders, R.J. (1997). Evolution of friendship and best friendship choices. In P. Doreian, F.N. Stokman, (Eds.), *Evolution of Social Networks* (pp. 149–65). Amsterdam: Gordan & Breach.

Leskovec, J., Adamic, L.A., & Huberman, B.A. (2007). The Dynamics of Viral Marketing. ACM, *Transactions on Web* 1, 1.

Leskovec, J., Backstrom, L., Kumar, R., & Tomkins, A. (2008). Microscopic evolution of social networks. *KDD'08,* August 24–27, 2008, Las Vegas, Nevada, USA. ACM 978-1-60558-193-4/08/08.

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review*, 3, 379–386.

Leventhal, H., & Scherer, K. (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition & Emotion*, 1, 3-28.

Lewicki, R.J., Bunker, B.B. (1996). Developing and maintaining trust in work relationships. In: Kramer, R.M., Tyler, T.R., (Eds.,) *Trust in Organizations: Frontiers of Theory and Research*. Sage Publications, Thousand Oaks, pp. 114–139.

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks,* 30: 330–342.

Li, X., & Chen, W. (2014). Facebook or Renren? A comparative study of social networking site use and social capital among Chinese international students in the United States. *Computers in Human Behavior*, 35, 116–123.

Lin, D. (1998). An information-theoretic definition of similarity. Proceedings of the *Fifteenth International Conference on Machine Learning*, 296 – 304.

Lin, H., Qiu, L. (2012). Sharing Emotion on Facebook: Network Size, Density, and Individual Motivation. *CHI 2012*, May 5–10, Austin, Texas, USA.

Lin, N., Dayton, P., & Greenwald, P. (1978). Analyzing the Instrumental Use of Relations in the Context of Social Structure. *Sociological Methods and Research* 7(2):1 49- 166.

Lin, N., Ensel, W. M., Vaughn, J.C. (1981). Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment. *American Sociological Review,46*(4), 393–405.

Lin, N. (1999). Building a Network Theory of Social Capital. *Connections* 22(1):28-51, INSNA.

Lin, C-J, Kuo, T-T, & Lin, S-D. (2014) .A Content-Based Matrix Factorization Model for Recipe Recommendation. *Springer*, Computer Science, Vol. 8444, pp 560-571.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. IEEE *Internet Comput.*, 7:76–80.

Lindquist, K.A., Wagr, T.D., Kober, H., Bliss- Moreau, E., & Barret, L.F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioural and Brain Sciences*, 35, 121–202.

Liu, N.-H.; Lai, S.-W.; Chen, C.-Y.; & Hsieh, S.-J. (2009). Adaptive music recommendation based on user behavior in time slot. *Int. J. Comput. Sci. Netw. Secur.*, 9, 219–227.

Long, J., Cunningham, F., & Braithwaite, J. (2013). Bridges, brokers and boundary spanners in collaborative networks: a systematic review. *Health Services Research*, 13:158, from http://www.biomedcentral.com/1472-6963/13/158.

Lops, P., Gemmis, M., Semeraro, G., Musto, C., & Narducci, F., (2013). Content-based and collaborative techniques for tag recommendation: an empirical evaluation. *Journal of Intelligent Information Systems*, February, Vol. 40, Issue 1, pp 41-61.

Lü L., Medo M., Yeung C.H., Zhang Y.C., Zhang Z.K., & Zhou, T. (2012). Recommender systems. *Physics Reports*, 519: 1–49.

Ma, H., Zhou, D., Liu, C., Liu, M.R., & King, I. (2011). Recommender Systems with Social Regularization. *WSDM’11*, February 9–12, 2011, Hong Kong, China.

Manovich, L. (2001). *The Language of New Media*. (Eds) The Myth of Interactivity. MIT Press.

Mayer, R.E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52.

Maynard, M. (2003). From global to local: How Gillette’s SensorExcel accommodates to Japan. *Keio communication Review*, 25, 57-75.

McKinsey Global Institute (2011). Big data: The next frontier for innovation, competition, and productivity. From: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CBsQFjAA&url=http%3A%2F%2Fwww.mckinsey.com%2F~%2Fmedia%2FMcKinsey%2Fdotcom%2FInsights%2520and%2520pubs%2FMGI%2FResearch%2FTechnology%2520and%2520Innovation%2FBig%2520Data%2FMGI_big_data_full_report.ashx&ei=SeTIU-C7HIua1AXW9YHICA&usg=AFQjCNEJyZHEjRELMRKoDqUgW1FR6PgxrA&sig2=RwlDebfiwK289Z9qPhIfLQ&bvm=bv.71198958,d.bGQ.

McLuhan, M. (1964). *Understanding Media, The Extensions of Man*. MIT Press, Cambridge, Mass.

Marmaros, D., & Sacerdote, B. (2006). How Do Friendships Form? *Quarterly Journal of Economics* 121:79-119.

Marsden, P.V. (2002). Egocentric and sociocentric measures of network centrality. *Social Networks*, Vol.24, no. 4, pp. 407–422, 2002.

Marsden, P.V., Karen, E., & Campbell, K.E. (1984). Measuring Tie Strength. *Social Forces*, Vol.63, No. 2, pp. 482-501.

Massa, P. & Avesani, P. (2007). Trust-aware Recommender Systems. *RecSys’07*, October 19–20, Minneapolis, Minnesota, USA.

McEvily, B., & Zaheer, A. (1999). Bridging ties: A source of firm heterogeneity in competitive capabilities. *Strategic Management Journal*, 20: 1133–1156.

McPherson, M., Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444.

McCaulley, M.H. (1999). The Myers-Briggs type indicator and leadership. In K. E. Clark & M. B. Clark, (Eds.), *Measures of leadership* (pp. 381–418).West Orange, NJ: Leadership Library of America.

McCroskey, J.C., Daly, J.A., Richmond, V.P., & Cox, B. (1975). The effects of communication apprehension on interpersonal attraction. *Human Communication Research*, 2, 51-65.

McCroskey, L.L., McCroskey, J.C., & Richmond, V.P. (2006). Analysis and Improvement of the Measurement of Interpersonal Attraction and Homophily. *Communication Quarterly*, Vol.54, Nº1, pp. 1-31.

McGraw, K.O., & Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1: 30-46.

Mead, G.H. (1934). M*ind Self and Society – From the Standpoint of a Social Behaviorist*. Chicago University Press, Chicago.

Meyer, W.U., Niepel, M., & Schutzwohl, A. (1994). Uberraschung und Attribution [Surprise and attribution]. In F. Forsterling & J. Stiensmeier-Pelster, (Eds.), *Attributions theorie. Grundlagen und Anwendungen* (pp. 105-121). Gottingen: Hogrefe.

Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychology*, 4, 59–74.

Messing, S., & Westwood, S.J. (2013). Friends that Matter: How Social Influence Affects Selection in Social Media. From http://www.stanford.edu/~seanjw/papers/POQTieStrengthNOTAnon.pdf.

Milgram, S. (1967). The Small-World Problem. *Psychology Today*, vol. 1, pp. 61-67.

Miller, K. (2011). *Organizational Communication: Approaches and Processes*. Cengage Learning; 6 edition.

Miranda, S.M., & Saunders, C.S. (2003). The Social Construction Of Meaning: An Alternative Perspective On Information Sharing. *Information Systems Research*, Vol.0, No.0, Month 2002, pp. 000–000.

Mislove, A., Gummadi, K.P., & Druschel, P. (2006). Exploiting Social Networks for Internet Search. Proceedings of the *7th Usenix/ACM SIGCOMM Internet Measurement Conference (IMC)*, San Diego, CA, from http://www.mpi-sws.org/~gummadi/papers/imc2007-mislove.pdf.

Mislove, A., Marcon, M., Gummadi, K.P., & Druschel, P. (2007). Measurement and analysis of online social networks. *IMC'07, SIGCOMM conference*, San Diego, California, USA.

Mislove, A., Viswanath, B., Gummadi, K.P., & Druschel, P. (2010). You Are Who You Know: Inferring User Profiles in Online Social Networks. *WSDM'10,* February 4–6, New York City, New York, USA. ACM, 978-1-60558-889-6/10/02.

Mizruchi MS. 2004. Berle and Means revisited: the governance and power of large US corporations. *Theory and Society*, 33:579–617.

Mollica, K.A., Gray, B., Trevino, L.K. (2003). Racial homophily and its persistence in newcomers' social networks. *Organization Science*. 14:123–36.

Monge, P., & Contractor, N. (2000). Emergence of communication networks. F. Jablin, L. Putnam, (Eds.), *Second Handbook of Organizational Communication*. Sage Publications, Thousand Oaks, CA.

Monge, P. R., & Contractor, N. (2003). *Theories of communication networks*. New York: Oxford University Press.

Moody, J. (2004). The structure of a social science collaboration network: disciplinary cohesion from 1963 to 1999. *American Sociology Review*, 69:213–38.

Mooney. R.J., & Roy, L. (2000). Content-based book recommending using learning for text categorization. In Proceedings of the *Fifth ACM Conference on Digital Libraries*, pages 195-204.

Mori, J., Sugiyama, T., & Matsuo, Y. (2005). Real-world Oriented Information Sharing Using Social Networks. *GROUP'05*, November 6-9, Sanibel Island, Florida, USA.

Myers, I.B. & McCaulley, M.H. (1985). *Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

Nanda, S., & Kotz, D. (2008). Localized Bridging Centrality for Distributed Network Analysis.in Proc. of *IEEE International Conference on Computer Communications and Networks - ICCCN*, Aug. 2008, pp.1-6.

Neumann, R., & Strack, S. (2000). “Mood Contagion": The Automatic Transfer of Mood Between Persons. *Journal of Personality and Social Psychology*. Vol.79, No.2, 211-223.

Neuman, Y. (2006). A theory of meaning. *Information Sciences*, 176 (2006) 1435–1449.

Nguyen, T.T., Hui, P-M, Harper, F.M., Terveen, L., & Konstan, J.A. (2014). Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity. *WWW'14*, April 7–11, Seoul, Korea.

Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 1977; 84:231–259.

Nooteboom, B., 1992. Towards a dynamic theory of transactions. *Journal of Evolutionary Economics* 2, 281–299.

Nooteboom, B. (2000). *Learning and Innovation in Organizations and Economies*. Oxford University Press, Oxford.

Nooteboom, B., Vanhaverbeke, W., Duysters, G.M., Gilsing, V.A., & van den Oord, A. (2005). Optimal cognitive distance and absorptive capacity. *ECIS* working paper 06-01.

Nooteboom, B., Haverbeke, W.V., Duysters, G., Gilsing, V., & Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy,* Nr. 36, 1016–1034.

Novak, T.P., Hoffman, D.L., & Yung, Y-F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science*, 19(1), 22–42.

Novak P.T., Hoffman, D.L., & Duhachek, A. (2003). The Influence of Goal-Directed and Experiential Activities on Online Flow Experiences. *Journal of Consumer Psychology*, 13(1&2), 3–16.

Novak, T.P., & Hoffman, D. L. (2009). The fit of thinking style and situation: New measures of situation-specific experiential and rational cognition. *Journal of Consumer Research*, 36, 56–72.

Oatley, K., & Jenkins, J.M. (1996). *Understanding emotions*. Cambridge, MA: Blackwell.

Oberg, A., & Walgenbach, P. (2008). Hierarchical structures of communication in a network organization. *Scandinavian Journal of Management*, 24, 183–198.

Obstfeld, D. (2005). Social Networks, the Tertius Iungens Orientation, and Involvement in Innovation. *Administrative Science Quarterly*, 50:100–130.

Ocepek, U., Bosnic, Z., Serbec, I.N., & Rugelj, J. (2013). Exploring the relation between learning style models and preferred multimedia types. *Computers & Education*, 69, 343–355.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245-251.

Ozsoylev, H.N., & Walden, J. (2011). Asset pricing in large information networks. *Journal of Economic Theory*, Vol. 146, Issue 6, November, Pages 2252–2280.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The pagerank citation ranking: Bringing order to the web. *Technical report*, Stanford Digital Library Technologies Project.

Pancke, S., & Gaiser, B. (2009). “With My Head Up in the Clouds": Using Social Tagging to Organize Knowledge. *Journal of Business and Technical Communication* 23, 318-349.

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review*, 3, 387–396.

Park, D.E., Kim, H.K., Choi, I.Y., & Kim J.K., (2012). A literature review and classification of recommender systems research. *Elsevier*. Expert Systems with Applications, Vol. 39, Issue 11, 1 September, Pages 10059–10072.

Paruchuri, S. (2010). Intraorganizational Networks, Interorganizational Networks, and the Impact of Central Inventors: A Longitudinal Study of Pharmaceutical Firms. *Organization Science*, Vol.21, No.1, pp. 63–80.

Pazzani, M.J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review*, 13: 393–408.

Perkins, D.N. (1981). *The mind's best work*. Cambridge, MA: Harvard University Press.

Perry-Smith, J.E. (2006). Social yet creative: The role of social relationships in facilitating individual creativity. *Acad. Management J.*, 49(1) 85–101.

Petróczi,A., Nepusz, T., & Bazsó, F., (2007). Measuring tie-strength in virtual social networks. *Connections*, 27(2): 39-52, INSNA.

Petrosky, D.G. (2011). *Relevance of Pre-web Social Network Theory to the Practice of Social Media Public Relations*. Master Thesis Dissertation, University of Rhode Island, USA.

Pink, D.H. (2008). *A Nova Inteligência*. Oficina do Livro, Lisboa.

Piscitelli, A., Andaime, I., & Binder, I. (2010). *El proyecto Facebook y la posuniversidade. Sistemas Operativos Sociales y Entornos Abiertos de Aprendizaje*. Ariel, Fundación Vodafone, Barcelona.

Ploghaus, A., Tracey, I., Clare, S., Gati, J. S., Rawlins, J. N. P., & Matthews, P. M. (2000). Learning about pain: the neural substrate of the prediction error for aversive events. Proceedings of the *National Academy of Science USA*, 97, 9281–9286.

Pollock, J.T. (2009). *Semantic Web for Dummies*. Wiley Publishing, Inc., Indianapolis, Indiana.

Pool, I. (1980). Comment on Mark Granovetter's 'The Strength of Weak Ties: A Network Theory Revisited.' *Meeting of the International Communications Association*, Acapulco, Mexico.

Press, L. (1995). McLuhan Meets the Net. *ACM*, Vol. 38, No. 6.

Ramiller, N.C., & Wagner, E.L. (2009). The element of surprise: appreciating the unexpected in (and through) actor networks. *Information Technology & People*, Vol.22 No.1, pp. 36-50.

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. Nature Reviews. *Neuroscience*, 4(3), 193–202.

Rangaswamy, A., Giles, C.L., & Seres, S. (2009). A strategic perspective on search engines: Thought candies for practitioners and researchers. *Journal of Interactive Marketing*, 23, 49–60.

Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. *Administrative Science Quarterly*, 48:240–67.

Reisenzein, R., Meyer, W., & Schützwohl, A. (1996). Reactions to Surprising Events: A Paradigm for Emotion Research. Proceedings of the *9th conference of the International Society for Research on Emotions* (pp. 292-296). Toronto: ISRE.

Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: the case of surprise. *Cognition and Emotion*, 14, 1–38.

Rentfrow, P.J., & Gosling, S.D. (2003). The Do Re Mi’s of Everyday Life: The Structure and Personality Correlates of Music Preferences. *Journal of Personality and Social Psychology*, Vol.84, No.6, 1236–1256, American Psychological Association.

Rivera, M.T., Soderstrom, S.B., & Uzzi, B. (2010). Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms. *Annual Review of Sociology*, 36:91-115.

Rogers, E.M., & Bhowmik, D.K. (1970). Homophily-Heterophily: relational concepts for communication research. Oxford Journals, Social Sciences, *Public Opinion, Quarterly*, vol. 34, Issue 4, pp. 523-538.

Rogers, E.M., & Kincaid, L.D. (1981). *Communication networks: Toward a new paradigm for research*. New York: Free Press.

Rogers, E. M. (1986). *Communication Technology: The New Media in Society*. New York: Free Press.

Rogers, E.M. (1994). *A history of communication study: A biographical approach*. New York: Free Press.

Rosen, D. (2009). Productivity and Performance in Academic Networks: Applications of Liaison Communication to Simmelian Ties, Structural Holes, and Degree Centrality. *Connections*, 29.2 (2009): 32-43.

Ruef, M. (2002). String ties, weak ties and islands: structural and cultural predictors of organizational innovation. *Industrial and Corporate Change*, Vol.11, Nº3, pp. 427-449.

Ruef, M., Aldrich, H., Carter N. (2003). The structure of founding teams: homophily, strong ties and isolation among U.S. entrepreneurs. *American Sociology Review*, 68:195–222.

Sarr, I., Missaoui, R., & Lalande, R. (2012). Dealing with Disappearance of an Actor Set in Social Networks. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. From: file:///C:/Users/Carlos%20Figueiredo/Downloads/SML.pdf

Scarantino, A. (2005). *Explicating Emotions*. Doctoral dissertation. University of Pittsburgh, Faculty of Arts and Sciences. USA.

Schafer, M.H. (2011). *Does Health Divide? Social Networks and Emergent Social Boundaries in a Retirement Community*. Thesis Dissertation, Purdue University, West Lafayette, Indiana, United States.

Schafer, J.B., Konstan, J.A., & Riedl, J. (2001). E-commerce recommendation applications. *Data Min Knowl Disc*, 5:115–153.

Schafer, J. B., Konstan, J., & Riedl, J. (1999). Recommender Systems in E-Commerce. In: EC ’99: Proceedings of the *First ACM Conference on Electronic Commerce*, Denver, CO, pp. 158-166.

Schmidt, L.A. (2010). Crowdsourcing for Human Subjects Research. *CrowdConf 2010*, October 4, 2010, San Francisco, CA.

Schein, A.I., Popescul, A., Ungar, L.H., & Pennock, D.M. (2002). Methods and metrics for coldstart recommendations. In: Proceedings of the *25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2002, 253-260.

Schutz, A. (1997). *Phenomenology of the Social World*. Northwestern University Press, Evanston, IL [1967].

Schutzwohl, A. (1998). Surprise and Schema Strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol.24, No.5, 1182-1199.

Schwarzer, R., & Knoll, N. (2003). Positive Coping: Mastering Demands and Searching for Meaning. in Lopez, S. J. & Snyder, C. R., (Eds.), *Handbook of Positive Psychological Assessment*. Washington, DC: American Psychological Association.

Seyerlehner, K (2010). *Content-Based Music Recommender Systems: Beyond simple Frame-Level Audio Similarity*. Doctoral dissertation, Johannes kepler University, Austria.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 1948; 27: 379–423. 623–656.

Shen, X., Tan, B., & Zhai, C. (2005). Implicit user modeling for personalized search. In Proceedings of *CIKM*, pages 824–831.

Shi, Y., Larson, M., & Hanjalic, A. (2014). Collaborative Filtering beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges. *ACM Computing Surveys (CSUR)*, Vol. 47, Issue 1, May, No. 3.

Shi, Z., Rui, H., & Whinston, A.B. (2013). Content Sharing in a Social Broadcasting Environment: Evidence from Twitter. From http://ssrn.com/abstract=2341243.

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin*, 86, 2, 420-428.

Siegel, D. (1999). *The Developing Mind: Towards a Neurobiology of Interpersonal Experience*. New York: Guilford.

Sih, A., Hanser, S.F., & A. McHugh, K.A. (2009) .Social network theory: new insights and issues for behavioral ecologists. *Behavioral Ecology and Sociobiology*, May, Vol. 63, Issue 7, pp 975-988.

Simmel, G. (1908). *The Sociology of Georg Simmel*. Free Press, New York

Sinha, R., & Swearingen, K. (2001). Comparing recommendations made by online systems and friends. In: Proceedings of the *DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*.

Smith, C.A., & Ellsworth, P.C. (1987). Patterns of appraisal and emotion related to taking an exam. *Journal of Personality & Social Psychology*, 52, 475-488.

Song, Y., Wang, H., & He, X., (2014). Adapting deep RankNet for personalized search. *WSDM '14*, Proceedings of the *7th ACM international conference on Web search and data mining*, Pages 83-92.

Sosa, M. (2011). Where Do Creative Interactions Come From? The Role of Tie Content and Social Networks. *Organization Science*, Vol.22, No.1, pp. 1–21.

Sommer, K.L., Baumeister, R.F., & Stillman, T.F. (1998). The Construction of Meaning from Life Events: Empirical Studies of Personal Narratives. In P. T. Wong & P. Fry, (Eds.), *The human quest for meaning* (pp. 143-161). Mahwah, NJ: Erlbaum.

Sparck Jones, K. (1988). A look back and a look forward. In: SIGIR ’88. Proceedings of the *11th Annual ACM SIGIR International Conference on Research and Development in Information Retrieval* (pp. 13-29). New York: *ACM*.

Sperber, D., & Wilson, D. (1995). Relevance: Communication and cognition (2nd ed.). *Padstow*, UK: T. J. Press.

Sternberg, R.J. (1998). A three facet model of creativity. In R. J. Sternberg, (Ed.), *The nature of creativity* (pp. 125–147). Cambridge: Cambridge University Press.

Stiensmeier-Pelster, J., Martini, A., & Reisenzein, R. (1995). The Role of Surprise in the Attribution Process. *Cognition and emotion*, 9 (1), 5-31.

Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., & Friston, K.J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Networks*, 18, 225–230.

Staber, U. (2004). Networking Beyond Organizational Boundaries: The Case of Project Organizations. *Creativity and Innovation Management*, Vol.3, Nº1, Blackwell Publishing, Ltd.

Stahl, G. A. (2000). A model of collaborative knowledge building. Paper presented at the *4th International Conference of the Learning Sciences*, Ann Arbor, MI.

Strange, B.A., & Dolan, R.J. (2001). Adaptive anterior hippocampal responses to oddball stimuli. *Hippocampus*, 11, 690–698.

Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., & Friston, K.J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw*., 18(3): 225–230.

Subrahmanyam, K., Reich, S. M., Waechter, N., & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults. *Journal of Applied Developmental Psychology*, 29(6), 420–433.

Sunstein, C.R. (2001). *Republic.com*. Princeton University Press, USA.

Sunstein, C.R. (2009). *Republic.com 2.0*. Princeton University Press, USA.

Susskind, A.M., Miller, V.D., & Johnson, J.D. (1998). Downsizing and Structural Holes. *Communication Research*, 25, 1, 30-65.

Swearingen, K. & Sinha, R. (2001). Beyond Algorithms: An HCI Perspective on Recommender Systems. *ACM SIGIR 2001 Workshop on Recommender Systems*.

Tadic, B., Gligorijevic, V., Mitrovic, M, & Suvakov, M. (2013). Co-Evolutionary Mechanisms of Emotional Bursts in Online Social Dynamics and Networks. *Entropy*, 15, 5084-5120.

Tan, B., Shen, X., & Zhai, C., (2006). Mining long-term search history to improve search accuracy. In Proceedings of *KDD*, pages 718–723, *ACM* Press, 2006.

Tang, J., Zhang, Y., Sun, J., Rao, J., Yu, W., Chen, Y., & Fong, A.C.M. (2012). Quantitative Study of Individual Emotional States in Social Networks. *IEEE Transactions on Affective Computing*, Vol.3 no.2, pp. 132-144.

Tavakolifard, M., & Almeroth, K.C. (2012). Social Computing: An Intersection of Recommender Systems, Trust/Reputation Systems, and Social Networks. *IEEE Network*, July-August, 0890-8044.

Teigen, K.H., & Keren, G. (2003). Surprises: low probabilities or high contrasts? *Cognition*, 87 (2003), 55–71.

Tewari, A.S., Kumar, A., & Barman, A.G. (2014). Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining. Advance Computing Conference (IACC), *IEEE International*, 500 - 503.

Tillema, T., Dijst, M., & Schwanen, T. (2010). Face-to-face and electronic communications in maintaining social networks: the influence of geographical and relational distance and of information content. *New Media Society*, 12: 965.

Tintarev, N., & Mastho , J. (2006). Similarity for News Recommender Systems. In Proc. of the *4th Int. Conf. of Adaptive Hypermedia and Adaptive Web-Based Systems* (AH-06).

Totterdell, P., Wall, T., Holman, D., & Epitropaki, O. (2004). Affect Networks: A Structural Analysis of the Relationship Between Work Ties and Job-Related Affect. *Journal of Applied Psychology*, Vol.89, No.5, 854–867.

Totterdell., P., Niven, K., Holman, D. (2010). Our emotional neighbourhoods. *The Psychologist*, Vol.23, nº 6, June. From http://www.thepsychologist.org.uk/archive/archive_home.cfm?volumeID=23&editionID=189&ArticleID=1681.

Tracy, J.L., & Randles, D., (2011). Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review*, Vol. 3, No. 4 (October), 397-405.

Trevino, L.K., & Webster, J. (1992). Flow in computer-mediated communication. *Communication Research*, 19(5), 539–573.

Tribus, M. (1961). *Thermostatics and Thermodynamics*. Van Nostrand, Princeton, New York.

Tversky, A., Kahneman, D. (1983). Probability, representativeness, and the conjunction fallacy. *Psychological Review*, 90 (4), 293–315.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 42, 35-67.

Uzzi, B., Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *The American Journal of Sociology*, Vol.111, No.2, pp. 447-504.

Vachon, M.L. (1982). Correlates of enduring distress patterns following bereavement: Social network, life situation and personality. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences*, 12, 783–788.

Valente, T.W., & Fujimoto, K. (2010). Bridging: Locating critical connectors in a network. *Social Networks*, 32 (2010) 212–220.

Valéry, P. (1957). In J. Hytier, (Eds.), *Oeuvres*, Gallimard, Paris.

Valenzuela, S., Park, N. & Kee, K.F. (2009). Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust and participation. *Journal of Computer-Mediated Communication*, 14, 875-901.

Van Duijn, M.J. van, Zeggelink, E.H., Huisman M, Stokman F.N., Wasseur, F.W. (2003). Evolution of sociology freshmen into a friendship network. J. Math. *Sociol*. 27:153–91.

Vargas, S., & Castells, P. (2011). Rank and Relevance in Novelty and Diversity Metrics for Recommender Systems. *RecSys'11*, October 23–27, 2011, Chicago, Illinois, USA.

Vinitzky, G., & Mazursky, D. (2011). The Effects of Cognitive Thinking Style and Ambient Scent on Online Consumer Approach Behavior, Experience Approach Behavior, and Search Motivation. *Psychology & Marketing*, Vol.28(5): 496–519.

Vinitzky, G., Mazursky, D. (2013). The Effects of Cognitive Thinking Style and Ambient Scent on Online Consumer Approach Behavior, Experience Approach Behavior, and Search Motivation. *Psychology & Marketing*, Vol. 28(5): 496–519, May.

Vladuțescu, S. (2012). Relationships and communication networks. *Journal of Community Positive Practices*, issue: 4, pages: 790796.

Vogel-Walcutt, J. J., Gebrim, J.B., Bowers, C., Carper, T.M., & Nicholson, D. (2011). Cognitive load theory vs. constructivist approaches: which best leads to efficient, deep learning? *Journal of Computer Assisted Learning*, 27(2), 133–145.

Waldstrøm, C. (2001). Informal Networks in Organizations – A literature review. From: http://pure.au.dk//portal/files/32302046/0003088.pdf.

Wang, H., & Wellman, B. (2010). Social Connectivity in America: Changes in Adult Friendship Network Size from 2012 to 2010. *American Behavioral Scientist*, 53(8) 1148–1169.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press, New York.

Watts, D. (2004). The “New” science of networks. Annual Review of Sociology, vol. 30: pp. 243-270.

Watts, D. (2007). A twenty-first century science. *Nature*, vol. 445, nr. 7127, pp. 489–489.

Watzlavick, P., Beavin, J. & Jackson, D. (1967). *Pragmatics of human communication*. New York: Norton.

Weber, M. (1947). The theory of social and economic organization. In A. H. Henderson & T. Parsons, (Eds.), *Glencoe*. IL: Free Press.

Weinberger, D. (2007). *Everything is Miscellaneous: The Power of the New Digital Disorder*. Holt Paperbacks. First Edition (April 29, 2008).

Wellman, B., & Wortley, S. (1990). Different Strokes from Different Folks: Community Ties and Social Support. *The American Journal of Sociology*, 96(3), 558–588.

Westphal, J.D., Milton, L.P. (2000). How experience and network ties affect the influence of demographic minorities on corporate boards. *Administrative Science Quartly*, 45:366–98.

White, H.C., Boorman, S.S., & Breiger, R.L. (1976). Social structure from multiple networks: I. Block-models of roles and positions. *American Journal of Sociology*, 81, 730-780.

Whitfield, J. (2008). *The secret of happiness: grinning on the internet*. In *Nature*, from http://www.nature.com/news/2008/080626/full/news.2008.918.html.

Williamson, K. (1998). Discovered by Chance: The Role of Incidental Information Acquisition in an Ecological Model of Information Use. *Library & Information Science Research*, Vol.20, Nº1, Pages 23-40.

Wimmer, A., & Lewis, K. (2010). Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook. *AJS*, Vol.116, Nº2: 583–642.

Woolf, B.P. (2009). *Building intelligent interactive tutors*. Burlington, USA: Elsevier Inc.

World Maps of Social Networks (2012). From: < http://vincos.it/world-map-of-social-networks/ >.

Wu, F., Huberman, B.A., Adamic, L.A., & Tyler, J.R. (2004). Information flow in social groups. From http://arXiv.org/abs/cond-mat/0305305.

Wuyts, S., Colombo, M.G., Dutta, S., & Nooteboom, B. (2005). Empirical test of optimal cognitive distance. *Journal of Economic Behavior and Organization*, 28, 277–302.

Yan, Y., Davison, R.M., & Mo, C. (2013). Employee creativity formation: The roles of knowledge seeking, knowledge contributing and flow experience in Web 2.0 virtual communities. *Computers in Human Behavior*, 29, 1923–1932.

Yan, S., Zheng, X., Chen, D., & Wang, Y. (2013). Exploiting two-faceted web of trust for enhanced-quality recommendations. Elsevier. *Expert Systems with Applications*. Vol. 40, Issue 17, 1 December, Pages 7080–7095.

Younus, A., Riordan, C., & Pasi, G. (2014). An Investigation into the Correlation between Willingness for Web Search Personalization and SNS Usage Patterns. From: ceur-ws.org/Vol-1127/paper11.pdf.

Zaman, M., Anandarajan, M., & Dai, Q. (2010). Experiencing flow with instant messaging and its facilitating role on creative behaviors. *Computers in Human Behavior*, 26(5), 1009–1018.

Zhang, L.F., & Sternberg, R. J. (2000). Are learning approaches and thinking styles related? A stud-y in two Chinese populations. *The Journal of Psychology*, 134, 469-489.

Zhang, L.F., & Sternberg, R. J. (2005). A threefold model of intellectual styles. *Educational Psychology Review*, 17(1), 1 – 53.

Zhang, L.F. (2009). Anxiety and thinking styles. *Personality and Individual Differences*, 47, 347-351.

Zhang, Y.C., Séaghdha, D., Quercia, D., & Jambor, T. (2012). Auralist: Introducing Serendipity into Music Recommendation. *WSDM'12*, February 8–12, Seattle, Washington, USA.

Zhou, J., Shin, S.J., Brass, D.J., Choi, J., & Zhang, Z. (2009). Social Networks, Personal Values, and Creativity: Evidence for Curvilinear and Interaction Effects. *Journal of Applied Psychology*, Vol.94(6), 1544-1552..

Zhou, T., Kuscsik, Z., Liu, J.-G., Medo, M., Wakeling, J.R., & Zhang, Y.-C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. Proceedings of the *National Academy of Sciences of the United States of America* 107, 4511-4515.

Zhou, T., & Lu, Y.B. (2011). Examining mobile instant messaging user loyalty from the perspectives of network externalities and flow experience. *Computers in Human Behavior*, 27(2), 883–889.

Ziegler, C.-N., McNee, S. M., Konstan, J. A., & Lausen, G. (2005). Improving recommendation lists through topic diversification. WWW 2005, May 10-14, Chiba, Japan.

APPENDICES

Appendix A – Online Questionnaire

The lineout presented here differs of the one seen online by participants, though, it contains the same contents than the original.

--------------------------------------- *1st screen* --

This research study is about social networks and information sharing.

Thank you for what you did on the previous phase (task 1).

For this second and last phase of the study (task 2), please read and answer to the questionnaire.

The questionnaire is formed by 5 groups, each one shown in one single screen. Please, be attentive when you scroll the screen for do not miss any question.

We expect that this survey will take about 15 minutes to complete.

In order to gain as much information as possible, please complete each question before moving on to the next.

Participation is entirely voluntary, and you may withdraw at any time by closing your browser window. Your results will be completely anonymous.

If you have any questions before completing this survey, please contact the researcher, Carlos Figueiredo (principle investigator) by e-mail: carlos.figueiredo@utexas.edu phone: 512-905-2414.

This research study has been reviewed by the Institutional Review Board for the Protection of Human Subjects the Human Research Protection Program at University of Texas at Austin. For information about the review process, please contact the (512) 232-2685 or the Office of Research Support at (512) 471-8871 or email: orsc@uts.cc.utexas.edu

IRB APPROVED ON: *04/09/2012* | IRB PROTOCOL # 2012-02-0141

Click to start the survey

-- *2nd screen* ---

- GROUP 1 -

This project is interested in people's social networks.

G1-1: At previous phase of this study you forwarded some posts of other participants to the message box of the project's Facebook page. You had selected those posts, because they stimulated an emotion in you. Given this, please write the names of the participants that posted the content in the project's Facebook page.

Please write the participants' names as they are registered on their Facebook pages. Please list up to 4 names. Start to list the names of those whose posts have stimulated the emotion of "surprise" on you.

All responses you provide will be kept strictly private and be used for the purposes of this study only.

Please write the names below also in a sheet of paper. This might help you to answer the questions that follow.

Person 1 [First name and initial of the last name] --

Person 2 [First name and initial of the last name] --

Person 3 [First name and initial of the last name] --

Person 4 [First name and initial of the last name] --

G1-2.1/ 2.2/ 2.3/ 2.4: The questions that follow are associated only with the names that you listed above. Please, follow the order that you established beginning by the first name listed.

First / 2nd / 3rd / 4th person that I listed above …	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Behaves like me (e.g. in public, among friends)	☐	☐	☐	☐	☐
Thinks like me (e.g. about life)................................	☐	☐	☐	☐	☐
Has similar interests	☐	☐	☐	☐	☐
Is different from me	☐	☐	☐	☐	☐
Expresses attitudes different from mine	☐	☐	☐	☐	☐
Has similar cultural heritage as I do .... (e.g. similar family traditions, behavior in public, likes the same media/music content)	☐	☐	☐	☐	☐

Click for next question

--- *3rd screen* ---

- GROUP 2 -

Now we would like to know more about how well you know the participants that you listed above.

G2-1: How often you had contact with each person that you mentioned above? Please select one option per person.

	Every day	Twice a week	Once a week	Twice a week	Twice a month	Twice a year	Once a year
1st person	☐	☐	☐	☐	☐	☐	☐
2nd person	☐	☐	☐	☐	☐	☐	☐
3rd person	☐	☐	☐	☐	☐	☐	☐
4th person	☐	☐	☐	☐	☐	☐	☐

G2-2: What type of relationship do you have with the people that you mentioned above? Please select the option following the order of your list of names. Please select only one option per row.

	Partner, Boyfriend/ Girlfriend	Direct family	Friend	Acquaintance	Other
1st person	☐	☐	☐	☐	☐
2nd person	☐	☐	☐	☐	☐
3rd person	☐	☐	☐	☐	☐
4th person	☐	☐	☐	☐	☐

G2-3: Indicate for how long you know each of the mentioned people.

	More than one year	More than three months	Less than three
1st person	☐	☐	☐
2nd person	☐	☐	☐
3rd person	☐	☐	☐
4th person	☐	☐	☐

G2-4.1/ 4.2: Please select the option if the statements match with the person listed.

Relatively to the persons listed above, I could ask to borrow a small sum of money to / I would contact I feeling sick, or needing health support:

	Yes	Uncertain	No
1st person	☐	☐	☐
2nd person	☐	☐	☐
3rd person	☐	☐	☐
4th person	☐	☐	☐

G2-5: On average, how close do you feel with the people that you listed at beginning of the questionnaire?

	Don't feel close at all	I don't feel very close	I feel reasonably close	I fell close	I feel very close
1st person	☐	☐	☐	☐	☐
2nd person	☐	☐	☐	☐	☐
3rd person	☐	☐	☐	☐	☐
4th person	☐	☐	☐	☐	☐

G2-6.1/ 6.2/ 6.3/ 6.4: Would you help the ***first / 2nd / 3rd / 4th person that you listed above…***

	Yes	Uncertain	No
Get information about a job	☐	☐	☐
Get information about a restaurant …..	☐	☐	☐
Get information about courses	☐	☐	☐
Provide emotional support	☐	☐	☐

G2-7: Who are the people that you mentioned above with whom you have private correspondence?

€ 1st person
€ 2nd person
€ 3rd person
€ 4th person

Click for next question

-- *4th screen* --

- GROUP 3 -

The following questions are about your musical preferences (genres) and emotions.

G3-1: Indicate your preferences by selecting a value between 1 (Very negative) to 10 (Very positive).

	1	2	3	4	5	6	7	8	9	10
Reflective and Complex (Blues, Folk, Classical, Jazz)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Intense and Rebellious (Alternative, Heavy metal, Rock)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Upbeat and Conventional (Country, Religious, Pop)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Energetic and Rhythmic (Funk, Hip-Hop, Soul, Electronica) ...	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

G3-2: Given the musical genres <u>above</u>, indicate the more common emotions when you listens music.

	Surprise (surprised, amazed, astonished)	Enjoyment (joyful, delighted, happy)	Interest	Distress (sad, downhearted, discouraged)	Anger (angry, mad, enraged)	Fear (afraid, scared, fearful)	Disgust (disgusted, feeling of distaste, feeling of revulsion)	Contempt (disdainful, contemptuous, scornful)	Shame	Guilt
Reflective and Complex .... (Blues, Folk, Classical, Jazz)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Intense and Rebellious .... (Alternative, Heavy metal, Rock)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Upbeat and Conventional .... (Country, Religious, Pop)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Energetic and Rhythmic .... (Funk, Hip-Hop, Soul, Electronica)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Click for next question

-- *5th screen* --

- GROUP 4 -

Please, write below your answers.

G4-3: Before you go to the last part of the survey, please write the names of the participants that you know best (people that you have been meeting socially/ professionally, e.g., in sports, parties, work, classes). Please write their names as they are registered on their Facebook pages. Do not list more than four names. However you can list less than four.

First name Last name

1- --- ---

2- --- ---

3- --- ---

4- --- ---

Click for next question

--- *6th screen* ---

- GROUP 5 –

Questions about Personal data

G5-1: **You are almost done.** We would briefly like to know a few things about you. What is your gender?

- € Male
- € Female

G5-2: How old are you?

G5-3: In order to organize your information, please tell us your:

First name

Last name

Email address

Name shown on your Facebook page

Residence (city, name of the building or dorm)

G5-4: What is your ethnicity?

- € Native American
- € African-American or of African descent
- € Middle Eastern
- € Asian-American or of Asian descent
- € Hispanic
- € White/ Caucasian
- € Other, please specify…

G5-5: What is your political affiliation (or your current strongest trend)?

€ Conservative
€ Moderate
€ Liberal
€ Independent
€ Other

G5-6: What is your current occupation?

€ Student
€ Employed
€ Self-employed
€ Unemployed
€ Retired
€ Other

G5-7: What kind of part-time job have you had in the last 2 years?

G5-8: What kind of activities (sports or hobbies) do you like the most?

G5-9: If you are a student, what is the highest education that you achieved / are currently attending?

€ Some High School coursework
€ High School
€ GED
€ Some College
€ Undergraduate Degree
€ Some Undergraduate Degree
€ Some Graduate School
€ Graduate Degree

G5-10: What is your religion spiritual practice?

€ Atheist
€ Catholic
€ Protestant
€ Christian
€ Jewish
€ Muslim
€ Hindu
€ Buddhist
€ Other, please specify

G5-11: What is your current marital status?

€ Single
€ In a Relationship
€ In an open Relationship
€ Engaged
€ Married
€ Divorced

G5-12: What is your family income?

- € Under $ 20,000
- € $ 20,000 - $ 40,000
- € $ 41,000 - $ 60,000
- € $ 61,000 - $ 80,000
- € $ 81,000 +
- € N/A

G5-13: And last but not least; **how many "friends" do you have in your Facebook?**

G5-14: In your Facebook, your friends have similar activities, common interests, or a similar general knowledge?

Please select the best option.

Majority of them are similar between each other

- € About 75% are similar between each other
- € About 50% are similar between each other
- € About 25% are similar between each other
- € Less than 25% are similar between each other
- € Just a few of them are similar between each other

--- *7th screen* ---

Thank you so much for your time and collaboration.

APPENDIX B – INFORMED CONSENT FORM

Title: Interpreting how homophily (similarities) and weak ties (social ties) intervene in the arousal of surprise in online social networks.

Introduction

You are being asked to participate in a research study. This form provides you with information about the study. The person performing the research will answer any of your questions. Read the information below and ask any questions you might have before deciding whether or not to take part. If you decide to be involved in this study, this form will be used to record your consent.

Purpose of the study

You have been asked to participate in a research study on how similarities between people and their social ties intervene in the process of information sharing in a social network.

What will you to be asked to do?

If you agree to participate in this study, you will be asked to complete two tasks:

- **Task 1:**
  - Share post, selected by you from your Facebook page with the project's Facebook page.
  - Forward the posts that stimulate in you an emotion to the message box of the project's Facebook page.
  - Write in the message the emotion that you perceived and the name of the participant that posted the content selected.
  - If the emotion was the "*surprise*", in few words, write why you were surprised.
- **Task 2:** Answer an online questionnaire.
  - You will receive an email with a link to access the online questionnaire.

The activities of this study will be spread by five days. In total you will spend about 1 hour to 1,5 hours. Task 1 will last four days. In total you will spend about 50 to 70 minutes (about 15 minutes per day). Task 2 will take 10 to 15 minutes. The study includes approximately 35-45 participants. There are no foreseeable risks in participating in this study.

You will receive a coupon for a brunch for completing the survey (task 1 and 2). Furthermore, for each new participant brought by you (up to a maximum of five), you will receive a coupon for a cookie.

In addition, you will be entered in a drawing for a chance to win either a Kindle Fire or one of three $20 Amazon gift certificates.
Participants that you bring to the study must send an email to carlos.figueiredo@utexas.edu for confirmation.

Participants must have an active Facebook account. Participants are invited to forward some content to the project's Facebook page. Are expected at least ten posts per participant. All participants will be "friends" in the project's Facebook page during the study.

To receive your compensation for completing the survey and for bringing additional participants please contact Carlos Figueiredo; T: (512) 905-2414; email: carlos.figueiredo@utexas.edu

Carlos Figueiredo; Ph.D. student of the University of Porto, visiting student of the University of Texas at Austin, TX 78712, Department of Radio-TV-Film, College of Communication.

Do you have to participate?

No, your participation is voluntary. You may decide to not participate at all or, if you start the study, you may withdraw at any time. Withdrawal or refusing to participate will not affect your relationship with The University of Texas at Austin in anyway.

If you would like to participate, you will receive a copy of this form.

What are my confidentiality or privacy protections when participating in this research study? This study is confidential. There is no way to connect your personal information with the interview data. Recordings will be kept for one year and then erased. The data resulting from your participation may be used for future research or may be made available to other researchers for research purposes not detailed within this consent form.

Whom to contact with questions about the study?

Prior to, during, or after your participation you can contact the researcher **Carlos Figueiredo** at **512-905-2414** or send an email to: carlos.figueiredo@utexas.edu

Whom to contact with questions concerning your rights as a research participant?

For questions about your rights or any dissatisfaction with any part of this study, you can contact, anonymously if you wish, the Institutional Review Board by phone at (512) 471-8871 or by email at orsc@uts.cc.utexas.edu.

Signature

You have been informed about this study's purpose, procedures, possible benefits and risks, and you have received a copy of this form. You have been given the opportunity to ask questions before you sign, and you have been told that you can ask other questions at any time. You voluntarily agree to participate in this study. By signing this form, you are not waiving any of your legal rights.

____________________ _________________

Printed Name Date

____________________ _________________

Signature of Person obtaining consent Date

APPENDIX C – INSTRUCTIONS TO STUDY PARTICIPANTS – PHASE 1

What will you to be asked to do *(please, repeat those actions during four days*)?

- **1:** *Friend* the project's Facebook page: Name: **Study Social-net Info-sharing** (Figure 2 below).
- **2:** *Share posts* from your Facebook with the project's Facebook page. Please, share at least **four posts** from your wall. Select the posts that are *meaningful to you* (Figure 3 below).
- **3**: *Forward the posts* that stimulate in you an emotion. Use the message box of the project's Facebook page to forward the posts (Figures 4 and 5).
    - Forward at least **five posts** registering your perceived emotion(s). If you want to forward more than five posts, even better. When forwarding a post, please include the following:
        I. Write the *emotion*(s) (see Figure 6). If the emotion was "*surprise*", in a few words, tells why you were surprised.
        II. Write the name of the *participant* that posted the content that you selected.
- **Note: It is not supposed comment the posts of other participants.**

Representation of the actions to be undertaken:

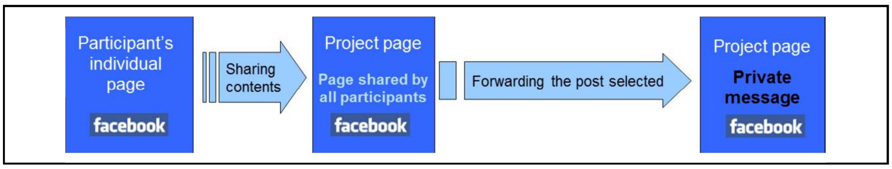


Figure 1 – Sharing content and message forwarding through project's Facebook page

Figure 2 – Find the page of the project and ask to *friend* the page.
Name of the project's Facebook page: **Study Social-net Info-sharing**

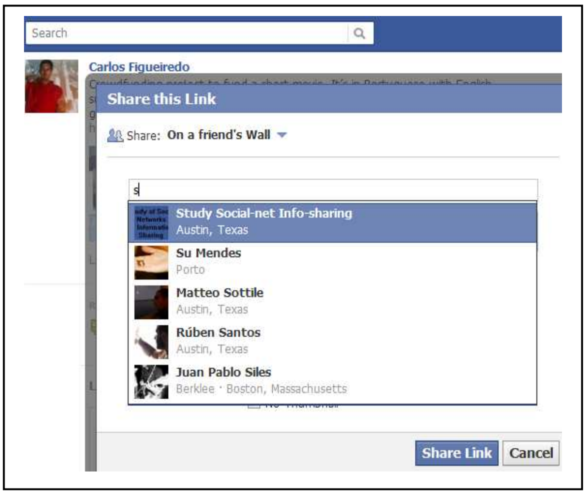


Figure 3 – Forwarding post from your wall to **Study Social-net Info-sharing.**

Figure 4 – Forwarding the post selected using the message box of the project's Facebook page.

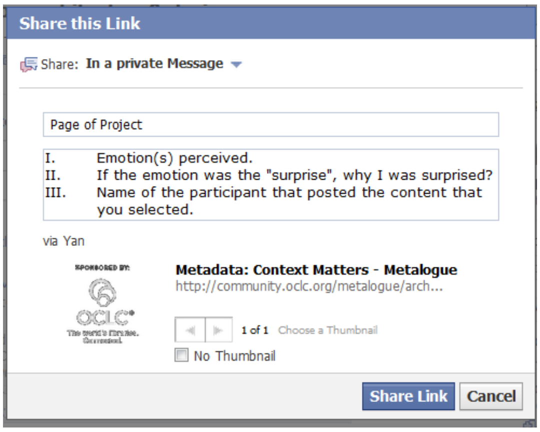


Figure 5 – Actions to do when forwarding the post selected.

To classify the emotions perceived select one or more options from the columns of Emotions and/ or Sub-categories. You can mix options to describe your emotion(s).

Emotions	Sub-categories
Surprise	surprised, amazed, astonished
Enjoyment	joyful, delighted, happy
Interest	---
Distress	sad, downhearted, discouraged
Anger	Angry, mad, enraged
Fear	Afraid, scared, fearful
Disgust	disgusted, feeling of distaste, feeling of revulsion
Contempt	disdainful, contemptuous, scornful
Shame	---
Guilt	---

Figure 6: Emotions Scale.

Note:

All participants need to be aware that the content shared must follow ethical rules about personal presence in social media, e.g. not sharing any kind of offensive, racist, xenophobic or pornographic content. Participants should not abuse their presence on the Facebook page of the project. Participants are invited to see more information online about the ethics of personal presentation on online social media.

E.g.,: http://research20atimperial.wordpress.com/compulsory-content/legal-ethical-issues/

Thank you very much for your collaboration.

Carlos Figueiredo, carlos.figueiredo@utexas.edu Phone: (512) 905-2414; University of Texas at Austin, Department of Radio-TV-Film, College of Communication. This research study has been reviewed by the Institutional Review Board for the Protection of Human Subjects the Human Research Protection Program at University of Texas at Austin. ***IRB APPROVED ON:*** *04/09/2012* | IRB PROTOCOL # 2012-02-0141.

Appendix D – Flyer

Hello, if you are a student **living in the Jester residence** this is for you.

I need of some volunteers for a short study on social network and information sharing. Participants need to be Jester residents, or be invited by participants that are Jester residents. Each participant will get a **FREE brunch** at Sagra Restaurant, a **FREE** Tiff's treat cookie for each additional friend invited, AND a **chance to win** an Amazon Kindle Fire or one of three $20 gift cards. Tasks include sharing Facebook posts and answering a questionnaire. Any help is greatly appreciated. Please, email me at carlos.figueiredo@utexas.edu saying that want to participate, and I'll email you with more information on how to participate. Thanks a lot!

Please, spread the word about the survey among your friends and acquaintances.
Your participation is greatly appreciated and you can make a difference.

Carlos Figueiredo, carlos.figueiredo@utexas.edu Phone: (512) 905-2414; University of Texas at Austin, Department of Radio-TV-Film, College of Communication. This research study has been reviewed by the Institutional Review Board for the Protection of Human Subjects the Human Research Protection Program at University of Texas at Austin. ***IRB APPROVED ON:*** *04/09/2012* | IRB PROTOCOL # 2012-02-0141