UNIVERSITÀ DEGLI STUDI DI TRENTO

CIMeC - Center for Mind/Brain Sciences

CiMeC
Center for Mind/Brain Sciences

Doctoral School in Cognitive and Brain Sciences
XXV Cycle

Doctoral Thesis

Collective Memories in Wikipedia

Michela Ferron

Advisor: Dr. Paolo Massa

December 2012

Abstract

Collective memories are precious resources for the society, because they contribute to strengthening the emotional bonding between community members, maintaining groups cohesion, and directing future behavior. Understanding the formation of the collective memories of emotional upheavals is important to a better comprehension of people's reactions and of the consequences on their psychological health. Previous studies investigated the effects of single traumatizing events, but few of them applied a quantitative approach to analyze the different psychological processes associated to the collective memories formation of upheavals on a large scale.

This thesis explores the opportunities of applying quantitative methods to the study of collective memories in a collaborative environment such as the English Wikipedia. First, the presence of commemoration processes in Wikipedia articles and talk pages about traumatic events will be investigated through the analysis of edit activity patterns. Second, natural language processing techniques will be applied to detect differences in the collective representations of traumatic and non traumatic events, in the temporal focus of old and recent traumatic events, and in the representations of natural and human-made disasters. Third, the temporal evolution of language related to emotional, cognitive and social processes will be analyzed in the talk pages of two different emotional upheavals, the 2005 London bombings and the 2011 Egyptian revolution. The results will confirm the interpretation of Wikipedia as a global memory place, and highlight specific psychological processes related to the formation of collective memories of different types of traumatic events, opening the way to the quantitative study of collective memory formation in digital collaborative environments.

Publications

Parts of this thesis (ideas, figures, results, and discussions) have appeared previously in the following publications:

Ferron, M. & Massa, P. (under review). Beyond the encyclopedia: Collective memories in Wikipedia. *Memory Studies*.

Ferron, M., & Massa, P. (2012). Psychological processes underlying Wikipedia representations of natural and manmade disasters. In *Proceedings of ACM WikiSym 2012 (August 27-29 2012)*. New York, ACM Press.

Ferron, M., & Massa, P. (2011). Collective memory building in Wikipedia: The case of North African uprisings. In *Proceedings of ACM WikiSym 2011 (October 3-5 2011)*. New York, ACM Press.

Ferron, M., & Massa, P. (2011). Studying collective memories in Wikipedia. Paper presented at the *3rd Digital Memories Conference, Prague, Czech Republic (March 14-16 2011)*.

Ferron, M., & Massa, P. (2011). Wiki Revolutions: Wikipedia as a lens for studying the real-time formation of collective memories of Revolutions. *International Journal of Communication*, *5*.

Ferron, M. (in press). Wikimemories: La costruzione della memoria collettiva del terremoto dell'Aquila in Wikipedia. In A. Micalizzi & M. Farinosi (Eds.), *Netquake. Media digitali e disastri naturali*. Milano: Franco Angeli.

Acknowledgements

This thesis could not have been completed without the precious support of many people, and I would like to express my sincere gratitude to them.

First and foremost, I want to thank my advisor Paolo Massa for doing a great job as supervisor and research colleague. During these years he provided me with his prompt support, encouragement, constructive criticism and friendship. He always found the time for listening to my doubts and providing enthusiastic guidance whenever I needed it. I am also very grateful to all the former members of the SoNet group at the Bruno Kessler Foundation, both for their ever prompt technical help and for creating such a friendly and pleasant working environment: Marco Frassoni, Maurizio Napolitano, Federico Scrinzi, Davide Setti and Asta Zelenkauskaite.

I am indebted to several actors whose support made this work possible. The LiveMemories project provided the financial support for this work and the opportunity to collaborate with experienced and competent researchers in a stimulating environment. In particular, I have benefited greatly from the advice and feedback on the statistical analyses given by Marco Baroni, from the valuable and constructive suggestions of Manuela Perrotta in the early phase of this project and from the later stimulating discussions with Alessandra Micalizzi.

I would like to express my special gratitude to my friends and family. In particular, I wish to thank Silvana, for her solid and warm help, Daniela, for being there when I needed it most, even during hard times, Giulia and Marta, for not giving up on me despite my discontinuous presence. I would like to express my gratitude to Giulio and Daniela, for their warmth and fondness during these years, to my mother, for her unconditional love and support, to my father, for opening up more and more, and to Lucia, for letting this happen. Finally, I would like to thank Alessandro. He is my strength, and there are no words to express it.

I dedicate this thesis to the memory of Ilenia Graziola, for the life that could have been and had not. I think of her at every achievement and every difficulty I encounter in my life.

Contents

Chapter 1

Introduction

On 11 September 2001, the US were shocked by one of the worst terrorist attacks in world's history: two hijacked airplanes were crashed into the Pentagon and hit the World Trade Center causing both towers to collapse, killing nearly 3,000 people. On 26 December 2004, a 9.1-magnitude earthquake struck off the west coast of Sumatra, Indonesia. The resulting tsunami devastated the coasts of fourteen countries with waves up to 30 meters high, killing 230,000 people. Few months later, four suicide bombers attacked the London's public transport system in the morning rush hour. It was 7 July 2005: the blasts killed 52 people and injured hundreds. Each of these shocking events left indelible marks on people's consciousness, producing widespread collective mourning, commemorative ceremonies and memorials (S. Brown & Hoskins, 2010).

Past research has found an increase in the degree of social sharing and rumination after traumatic events, when people try to find explanations and share their feelings and emotions with others (Mehl & Pennebaker, 2003; Pennebaker & Harber, 1993; Pyszczynski, Solomon, & Greenberg, 2003; Rimé, Finkenauer, Luminet, Zech, & Philippot, 1998; Rimé, Philippot, Boca, & Mesquita, 1992). This period of increased social sharing provides the basis for the formation of collective memories about these emotional upheavals (Pennebaker & Banasik, 1997). Through their normative and directive function, collective memories play an important social role, influencing not only attitudes toward the past, but also toward the present and the current society (Paez, Basabe, & Gonzalez, 1997; Wang, 2008). They define the identity of a community, maintaining group cohesion and strengthening emotional bonding between members (Wang, 2008). Persisting for years and generations, collective memories can be at the root of prejudice, nationalism and wars, directing future behavior, mobilizing collective action and leading to social and political change (Alexander, Eyerman, Giesen, Smelser, & Sztompka, 2004; Eyerman, 2004; Pennebaker, Paez, & Rimé, 1997; Wang, 2008).

Thus, understanding how they are formed through the public discourse is crucial to a better comprehension of the society (Irwin-Zarecka, 1994; Paez et al., 1997; Sturken, 1997). In the long term, examining the formation of collective memories can provide insights into how a culture functions, and how cultural arenas, political and social issues intersect (Sturken, 1997). In the short term, analyzing the social sharing in the immediate aftermath of emotional upheavals can allow a better comprehension of sensemaking processes, of typical responses to extremely stressful situations and of the effects of victimization (Baum, 1987).

However, few studies have attempted to empirically investigate on a large scale how communities form their collective memories, and how the psychological processes involved in the development of collective memories evolve into social

contexts. Past research investigated the psychological outcomes of traumatic events employing, by and large, interviews and retrospective self-reports, which can lead to biases and memory distortions (A. A. Stone et al., 2000). With the recent development of Internet and Web 2.0, researchers can now unobtrusively tap into large amounts of textual data about people's thoughts and reactions to emotional upheavals, exploiting new opportunities toward quantitative and longitudinal studies on a large scale. Studying words usage patterns in the context of traumatic events can provide valuable insights into the psychological processes taking place in the construction of collective memories.

The collaborative environment of the online encyclopedia Wikipedia is particularly appropriate for the study of emotional upheavals, because users actively participate in the construction of shared narratives, negotiating their perspectives to reach a common view of events (Ferron & Massa, 2011b; Keegan, 2011; Pentzold, 2009). Following Pentzold's (2009) interpretation of Wikipedia as a global memory place, where collective memories are built, the unifying theme of this dissertation is the investigation of collective memory building processes and the shared representation of traumatic events in the English Wikipedia. In the thesis, I provide evidence of the presence, in Wikipedia, of the typical processes characterizing collective memory building. In addition, I apply quantitative text analysis to study the psychological processes detectable in the collective representations of different traumatic events in Wikipedia, and their temporal evolution in users' discussions during the immediate aftermath of emotional upheavals.

The first contribution of this thesis is the exploitation of quantitative analysis of edit activity and of words usage patterns for the study of collective memories and the associated psychological processes. Specifically, I apply quantitative analysis of users' edit activity on articles and talk pages about traumatic events, showing the presence of commemoration patterns typical of collective memory building and maintenance processes, and hence empirically supporting the interpretation of Wikipedia as a global memory place. In addition, I apply quantitative text analysis to Wikipedia articles dedicated to traumatic events. I show that word count approaches can be a valuable tool to distinguish the diverse psychological processes detectable in the representation of traumatic and non traumatic events and of natural and human-made disasters. Although much of the past literature investigated the effects of single traumatizing events (Adler, 1943; Cohn, Mehl, & Pennebaker, 2004; Gleser, Green, & Winget, 1981; Gleser, Green, & Winget, 1978; Green, 1980; Henderson & Bostock, 1977; Leopold & Dillon, 1963; Ploeger, 1977; Titchener & Kapp, 1976), few studies compared natural and man-made disasters, because the specificity of each traumatic experience makes it difficult to outline common effects of different types of emotional upheavals. However, past research proposed theoretical reasons for hypothesizing substantial differences in the psychological consequences of natural and human-made disasters (Baum, 1987; Gleser et al., 1981). Drawing from the characterizing properties proposed by Baum (1987) and employing quantitative text analysis, I provide evidence of distinct psychological processes underlying the representation of natural and human-made traumatic events in Wikipedia articles, validating past findings and extending them to digital environments.

The second contribution is the study of the temporal evolution of psychological processes which can be derived from words usage patterns during the immediate aftermath of traumatic events. Previous research applying word count approaches analyzed the psychological responses of individuals to the September 11 attacks of

2001 focusing on blogposts and text pagers (Back, Küfner, & Egloff, 2010; Cohn et al., 2004), or studied the language used in Wikipedia articles about airplanes disasters (Keegan, 2011). However, an empirical investigation of how people's reactions evolve in a truly collaborative environment was missing. To this end, Wikipedia talk pages related to emotional upheavals are particularly appropriate because they are created with the specific purpose of allowing users to discuss and negotiate different perspectives in order to reach a shared narrative of the events. In this dissertation I present a study of the temporal evolution of affective, cognitive and social processes in Wikipedia talk pages during the immediate aftermath of the 2005 London bombings, and during the first days of the 2011 Egyptian revolution. The results will partially validate previous findings by Cohn and colleagues (2004), yet suggesting a different temporal evolution of affective and cognitive processes.

Summarizing, the main research questions that I pursue in this manuscript are: 1) Can Wikipedia be interpreted as a global memory place, and does it encompass the typical processes of collective memory building? 2) Are word count approaches effective in detecting differences in the Wikipedia representations of traumatic and non traumatic events, and in the temporal focus of old and recent emotional upheavals? 3) What are the psychological processes embedded in Wikipedia collective representations of natural and human-made disasters? 4) How do different psychological processes evolve over time during the first days following traumatic events in collaborative environments such as Wikipedia talk pages?

In order to answer these questions, the manuscript is organized as follows. In Chapter 2, I provide some background on past research on collective memories and on the online encyclopedia Wikipedia. Chapter 3 addresses the first research question analyzing edit activity patterns in articles and talk pages about traumatic events, investigating the presence in Wikipedia of the main collective memory processes outlined in past research. In Chapter 4, I discuss the advantages and limitations of word count approaches, applying Linguistic Inquiry and Word Count (Pennebaker, Francis, & Booth, 2001) to analyze the psychological processes detectable in articles about traumatic and non traumatic events, and the temporal focus in articles about old and recent emotional upheavals. Chapter 5 is devoted to analyze differences between natural and human-made disasters, focusing on distinct psychological processes identifiable in their representation in Wikipedia articles. In Chapter 6, I consider two specific events, both emotionally charged but substantially different, the 2005 London bombings and the 2011 Egyptian revolution, to analyze the temporal evolution of linguistic patterns expressing specific psychological processes in the collaborative environment of Wikipedia talk pages. Finally, Chapter 7 concludes the dissertation summarizing the main achievements of the thesis.

Chapter 2

Background

2.1 Collective memory and cultural trauma

Contemporary perspectives of remembering understand memory as an active process, where what is remembered is actively built and reconstructed every time, introducing the past into the present (see for example S. Brown & Hoskins, 2010; Garde-Hanse, Hoskins, & Reading, 2009). This active and dynamic role of memory can be dated back to the psychological tradition, and particularly to Bartlett (Bartlett, p. 213), who also argued that our memory is influenced by the presence of others and by our social organization:

> Remembering is not the re-excitation of innumerable fixed, lifeless and fragmentary traces. It is an imaginative reconstruction, or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience.

The term "collective memory" was initially introduced by Halbwachs (1992), and since then it has been used in different disciplines, such as sociology, history, philosophy and psychology, to refer to a wide range of phenomena. As developed in the Durkheimian sociological tradition by Halbwachs, individual memory and individual identity are always conceived and mediated by some collectivity. For instance, individual recollections depend on the settings and collectives in which they are expressed, such as workplaces or families (for example Middleton & Brown, 2005). In this perspective, memory is collective because it is supra-individual, in that it is located in the social resources that shaped it, and individual memory cannot be seen as detached from social factors and social influences. Indeed, even if it may be the individual who stores and recollects information, every step of the memory processes is embedded in the social environment and is influenced by the social resources provided by the environment (Eyerman, 2004; Halbwachs, 1950; Hirst & Manier, 2008).

In this work, following Olick and Levy (1997), collective memory is intended as a continuous active process of sense-making and negotiation between past and present. As Zerubavel (2003, p. 28) pointed out, collective memory is more than the combination of individual memories: to gather the social shape of the past we need to consider the "social time lines" built by mnemonic communities, thus investigating "social sites of memory". In fact, it has been argued that the relationship between individual and collective memory can be examined empirically, studying collective and social remembering (Hirst & Manier, 2008; Wertsch, 2002).

In the past, many different research traditions have contributed to the investigation of collective remembering processes. Several psychological studies

analyzed experimentally the processes involved in remembering in social groups using the collaborative recall paradigm (see for example Barnier & Sutton, 2008; Barnier, Sutton, Harris, & Wilson, 2008; Cuc, Ozuru, Manier, & Hirst, 2006; Ekeocha & Brennan, 2008; Harris, Paterson, & Kemp, 2008; Weldon & Bellinger, 1997). Typically, in a collaborative recall experiment a group of participants learn a set of stimuli, such as a words list or pictures, recall them individually or in group, and finally perform a last individual recall test. In order to obtain a measure of the effects of collaboration on the accuracy of recall, the number of items recalled by collaborative groups is compared with the number of items recalled by nominal groups. The latter is calculated as the sum of non-overlapping items recalled by the same number of participants forming a collaborative group. One of the most deeply studied aspects of group remembering in collaborative recall is collaborative inhibition (e.g., Weldon & Bellinger, 1997), a robust effect consisting in a lower performance of collaborative groups with respect to nominal groups. In other words, while groups unsurprisingly recall more than individuals alone, they fail to elicit new memories and to cross-cue each other, thus not performing at their full potential.

Other studies addressed the mechanisms underlying the formation and maintenance of flashbulb memories (R. Brown & Kulik, 1977; Conway, 1995; Winograd & Neisser, 1992). Flashbulb memories are vivid, detailed and long lasting memories people have about the circumstances in which they first learned about particularly shocking and emotionally charged events, such as the assassination of public figures. One of the most investigated aspects related to the formation and maintenance of flashbulb memories is the role of emotions. To this regard, several studies evidenced that higher levels of emotional arousal are associated with a better recall of the circumstances in which participants discovered the related events (e.g., Bohannon, 1988; Finkenauer, Gisle, & Luminet, 1997; Pillemer, 1984; Rubin & Kozin, 1984).

On the other hand, a different research tradition has examined the social and conflicting aspects of collective remembering. For instance, Zerubavel (2003), drawing on numerous examples ranging from the Holocaust to Hiroshima, from Native Americans to the conflict between Serbs and Albanians in Kosovo, highlighted the essentially social nature of memory. The author showed how complex and multifaceted is the process through which mnemonic communities remember the past, and how individuals and communities socially shape their collective memories, organizing the past and bridging historical events into coherent narratives in order to ensure its continuity.

Marita Sturken (1997) investigated the centrality of remembering in the American culture, analyzing how cultural memory is closely bound to political and social issues, often revealing conflicting arguments and divisions. Concentrating primarily on the Vietnam Veterans Memorial and the AIDS Memorial Quilt, Sturken investigated memory building focusing on "what its telling reveals about how the past affects the present" (1997, p. 2). Thus, her analysis explored the role of cultural memory processes, such as the reenactment of historical events, as means for healing and forgetting, interpreting the memorials as sites for mourning and grief.

As it also emerged from the works previously mentioned, Wertsch (2002) observed that while the psychological tradition have studied memory from an individual perspective, focusing mainly on its accuracy, sociologists and other scholars of collective memory have focused on more global and cultural aspects, such as political processes, active contestation and negotiation. In order to gain a

deeper understanding of collective memory processes the author called for the need to overcome this "functional dualism" (2002, p. 31), given by the implicit division between these strands of research. This can be done by recognizing that memory serves both of these functions, that is, on the one hand, to give an accurate representation of the past, and on the other hand to serve the needs of a community in the present, which are often connected to its collective identity, such as the need to mourn or to get over a past trauma.

With this regard, the online encyclopedia Wikipedia encompasses both of the aspects described by Wertsch because, while it aims at "representing fairly, proportionately, and as far as possible without bias, all significant views that have been published by reliable sources" ("Wikipedia:Neutral Point of View", 2001), it also pursues broader scopes, such as, as explained later in this dissertation, that of educating people, providing editors with the opportunity to share their feelings about traumatic events and strengthening their sense of belonging to the community by working together.

Indeed, Web 2.0 platforms and in particular Wikipedia offer a new opportunity for enriching collective memory research, opening the possibility for new forms of collaborative remembering. Moreover, digital archives allow social scientists to access huge amounts of data about people's thoughts, feelings and communicative interactions in an unobtrusive way and almost in real time.

One of the favorite perspectives for the study of collective memories processes is in the framework of cultural trauma research (Alexander et al., 2004). As Eyerman (2004) pointed out, trauma as a cultural process is closely connected with the formation of emergent collective memory and collective identity of a group. Unlike psychological trauma, which entails a prominent emotional wound for the individual, cultural trauma affects a social group with some degree of cohesion, and the trauma itself does not need to be experienced by every group member. Most importantly, even if it is possible to identify some particular event that triggered the traumatic sequence, collective trauma is socially mediated by a number of cultural processes. In fact, events are not traumatic by themselves, but they are attributed a traumatic meaning when they shake the foundation of a groups' collective identity (Alexander et al., 2004).

Still, even if cultural trauma is a construction of the society, it is possible to outline a number of features which characterize potentially traumatic events. According to Neal (Neal, 1998, pp. 9-10), the traumatizing event should be a "volcano-like", "extraordinary event" that causes "disruption" and "radical change … within a short period of time". Sztompka (2000) pointed out that to be potentially traumatizing, an event should be sudden, radical and deep, perceived as imposed from outside, unexpected, surprising, shocking and repulsive. Sztompka provided also a list of social changes that meet these characteristics, and are therefore in condition to trigger a cultural trauma. Among these, there are revolutions, genocides, deportations and ethnic cleansings, mass murders or assassinations of political leaders, acts of terrorism and lost wars.

These precise characterization allows to consider events such as the World Trade Center attacks of 11 September 2001, the tsunami that devastated Indonesia on 26 December 2004, the London bombings of 7 July 2005, or the Egyptian revolution started in 2011 and classify them as traumatic events. So, in order to reduce the ambiguity around what can be considered an object of collective memory processes

or not, this dissertation will focus on the formation of collective memories of traumatic events.

2.2 Web 2.0 and social network memories

In the last decades social memory studies grew substantially, producing a number of diverse conceptual definitions and typologies of collective memory (Harris et al., 2008; Hirst & Manier, 2008; Olick, 2008). Wertsch (2002) pointed out that, as human beings are essentially social in nature, remembering is situated in a particular cultural context and is mediated by the sociocultural tools provided by the society. In this way, remembering is also a distributed process which involves both the individuals who remember and the tools that serves the remembering process, such as narratives or, especially in the Internet era, computers. Garde-Hanse, Hoskins and Reading (2009) proposed that the influence of media and their associated technologies can be at least in part accounted for the recent increased and diffused attention to memory studies. Television has increasingly contributed to the formation of collective memories with movies, news, TV documentaries and commemorations (Sturken, 1997), for example with the mini-series on NBC in 1978 on Holocaust (Shandler, 1999).

Today, the widespread diffusion of new digital media and technologies raises the need for a paradigmatic shift that takes adequately into account the novel relationship between memory and media (Hoskins, 2009a). Our everyday life is indeed deeply meshed with digital technologies: we are overwhelmed by the information overload and the excess of connectivity in our lives, but at the same time we are always online, worried about information loss, obsessed with recording, keeping track, retrieving and archiving information. In this new "digital memory culture" (Hoskins, 2009b), we are provided with virtual places where we can edit our lives, store and make immediately accessible affective and personal moments on the Internet. We can share online personal memories, such as photographs, videos, text messages, blogs, digital archives and storytelling (Hoskins, 2009a). Through Web 2.0 and social networking sites (boyd & Ellison, 2007), we can contribute content as easily as we can consume it (Hoskins, 2009b). We can interact publicly on platforms such as Facebook or MySpace, talking about our lives and sharing our personal thoughts and emotions; we can upload family videos and photographs on YouTube and Flickr; and we can also collaboratively write the story of public events such as September 11 attacks or 7 July 2005 London bombings on Wikipedia.

In this sense, digital media allow the formation of a new hybrid form of memory. This new memory is public but also private in that although personal digital memories can be continuously modified, they are also resistant to total erasure (Hoskins, 2009b). New digital media and their associated technologies also provide fast access to sites of memory and collective identity in times of mourning and grief (Hoskins, 2009a). In other words, the contemporary bottom-up participatory culture allows for an evolution of memory production processes: memories are not anymore just consumed, but also creatively produced in a participatory and decentralized way. Thus, the resulting new memory, dynamically formed through our digital practices, our interactions with technology, and our connection to digital networks, is potentially visible and accessible to everyone. Moreover, it is fluid, because it can be collectively shaped and modified. This memory also becomes a social network

memory, which has the potential for forging and sustaining communal relationships and activities and, since it is based on peer-produced relationships enabled through social networking, it is "fluid, de-territorialised, diffused and highly revocable, but also immediate, accessible and contingent on the more dynamic schemata forged through emergent sociotechnical practices" (Hoskins 2009b, p. 29).

Interestingly, the widespread accessibility of these online digital practices and collective memory building processes allows for new research opportunities towards empirical and quantitative work on a large scale. Researchers can now directly access the communicative exchange between Web 2.0 platforms' users representing people's thoughts and feelings without giving up the spontaneity of interactions. The Internet provides a large amount of data which researchers can collect in an unobtrusive way and almost in real time, and the massive backup into digital archives allows researchers to conduct longitudinal studies on these data (Cohn et al., 2004).

2.3 The online encyclopedia Wikipedia

One of the most stimulating Web 2.0 environments for the study of collective memory building processes is Wikipedia. Launched in 2001, Wikipedia is an online encyclopedia built by the independent and voluntary work of millions of people, apparently without receiving explicit credit for their contributions (Forte & Bruckman, 2005). In fact, Wikipedia pages can be edited directly on the website by anyone. Pentzold (2009) proposed that the online encyclopedia can be seen as a global memory place, a concept drawn from the Pierre Nora's *lieu de mémoire* or "site of memory", which is where "memory crystallizes and secretes itself" (Nora, 1989, p. 7). This is especially true in the case of Wikipedia pages about traumatic events, such as for instance the September 11 attacks, where memory is negotiated through the discussion of different points of view (see *Figure 2.1*).

Figure 2.1 Screenshot of the article about September 11 attacks.

In Wikipedia, article pages constitute the main type of content. But in addition to articles, Wikipedia contains several other types of pages, which are used to discuss about the content of articles, debate on Wikipedia's policies and procedures, or directly communicate with other users. For instance, every Wikipedia article has an associated talk page in which anyone can freely discuss the article content and

structure, negotiate and suggest improvements or changes (see *Figure 2.2*). Usually contributions are signed and this allows to connect every edit to the user who wrote it. Anonymous users, who do not authenticate themselves with a personal login and a password, are identified by the IP address of the computer from which they are connected. Registered users, who are identified by a nickname, also have a user page on Wikipedia, which is a personal page where they can present themselves, and a user talk page for direct communication with other users.

Figure 2.2 Screenshot of a piece of discussion in the talk page associated to the article on September 11 attacks.

The Mediawiki open source web platform, powering Wikipedia, records any change made by any user to any article or talk page so that it is possible to reconstruct the history of each user's contributions (see *Figure 2.3*). Indeed, every edit is recorded along with the respective date and time, the author and the optional comment that justifies the edit.

Figure 2.3 Screenshot of the revision history page associated to the article on September 11 attacks.

At September 2012, the English Wikipedia counted more than 17 million users, about 28 million pages and more than 4 million articles, which makes Wikipedia an

interesting platform for empirical studies of social processes on a large scale ("Wikipedia:Size of Wikipedia", 2001).

In the last years, the increasing popularity of Wikipedia and the availability of large datasets have attracted a considerable amount of scientific research. For example, some academics have tried to quantify the growth of Wikipedia and unveil its network dynamics, finding that some of the structural properties of other network contexts, such as the exponential growth, the power-law degree distribution and the preferential attachment, are also found in the online encyclopedia (see for instance Voss, 2005). Other researchers have tried to assess the quality of content in Wikipedia. For example, Stvilia and colleagues (2005) used exploratory factor analysis to find metrics to measure article quality, built on several characteristics of articles, such as the number of anonymous edits, unique editors, or article length. Another research compared a set of 42 articles about scientific topics in Wikipedia and the Encyclopedia Britannica, finding minimal differences in accuracy, with an average of four errors for Wikipedia and three errors for Britannica (Giles, 2005).

Although most of the academic efforts concentrated on the actual content of the online encyclopedia focusing on article pages, most relevant to the study of collective memory formation in the online encyclopedia are works about collaboration and social aspects of Wikipedia. Viégas and colleagues (2007) carried on an empirical analysis of social aspects and collaboration mechanisms of Wikipedia to better understand how the community had evolved. Consistently with previous findings by Bryant, Forte, and Bruckman (2005), they observed that editing activity for coordination and administrative purposes grew faster than direct contributions to articles. Comparing results obtained in a previous study (Viégas, Wattenberg, & Dave, 2004), they found that between 2003 and 2005 the user talk and the Wikipedia guidelines were the namespaces with the highest rate of growth. Using manual coding to analyze a subset of 25 talk pages, they found that most of all, talk pages were used for requesting editing coordination, showing their crucial role in allowing users to ask for help and discuss editing activity in advance.

On the same direction, Kittur and colleagues (2007) studied the rise of conflict and the costs of coordination in Wikipedia focusing on "indirect work", defined as "excess work in the system that does not directly lead to new article content" (p. 453). They showed that, although remaining the dominant activity, direct work on articles decreased from 90% of edits in 2001 to 70% in 2006, while indirect work such as discussion, coordination and maintenance activity increased from 2% to 12%.

Wikipedia is also an excellent source of data for Social Network Analysis (SNA), since it provides large datasets in different languages, which researchers can access unobtrusively and in real time. Offering many different types of pages and several ways for users to interact, Wikipedia also provides a more complex social context compared to other online communication tools and platforms such as Internet Relay Chat, email or Twitter. The richness and the diversity of possible interactions in Wikipedia allow researchers to represent these relations in a variety of different network graphs.

Interesting for the study of social processes in Wikipedia are works about communication networks of user talk pages (UTPs), devoted to direct communication between users, and article talk pages (ATPs), dedicated to discussion about article content (see for instance Laniado, Tasso, Volkovich, & Kaltenbrunner, 2011; Massa, 2011; Viegas et al., 2007; Welser et al., 2011). Usually, in networks of

UTPs nodes are users and an edge is drawn from A to B when A leaves a comment on the user talk page of B. Networks of ATPs can be extracted exploiting their tree structure, where nodes are Wikipedia users and a directed edge links user A to B if A has written at least one comment indented under an entry by user B (Laniado et al., 2011). Welser and colleagues (2011) used ego networks (i.e., graphs centered around a specific user) to find patterns of structural signatures by analyzing the distribution of edits across different types of pages and relationships between editors. Similarly, extracting networks of direct replies between users in ATPs and UTPs, Laniado and colleagues (2011) studied structural patterns of interaction, focusing on directed assortativity, which is the tendency of nodes in a network to connect to other nodes with a similar number of edges.

Other works tried to improve the understanding of communication networks in Wikipedia by combining network and content analysis (Welser, Underwood, Cosley, Hansen, & Black, 2010) or adding attributes to edges, such as a binary classification into positive or negative interaction (Zhu, Kraut, Wang, & Kittur, 2011). Welser and colleagues (2010), using the network visualization tool NodeXL (Smith et al., 2010), proposed to combine network and content analysis to assess the quality of talk pages, visualizing discussion patterns, identifying high quality contributors and locating problematic interactions. Maniu, Cautis and Abdessalem (2011) proposed algorithms for extracting a signed network from Wikipedia edit activity considering edits over commonly authored articles, other activities such as votes for adminship, the restoring of an article to a previous version, or the assignment of barnstars (a prize acknowledging valuable contributions). Indeed, an edit on an article talk page or user talk page can represent alignment and positive feedback as well as disagreement and negative feedback. Zhu and colleagues (2011) exploited machine learning models to automatically identify four types of leadership behaviors (positive, negative, directive and social) from 4 million messages sent between Wikipedia users, finding that editors in peripheral roles produce a large proportion of leadership behaviors.

Another research topic on Wikipedia tried to analyze cross-cultural differences among versions of the online encyclopedia in different languages. Pfeil and colleagues (Pfeil, Zaphiris, & Ang, 2006) examined the evolution of the article "game" in the French, German, Japanese, and Dutch Wikipedia, finding that the cultural dimensions identified by Hofstede (1991) in an organizational context can also be observed in online environments. For instance, they found that countries with a high power distance (i.e., countries which tend to emphasize hierarchical power and are characterized by a large inequality in the distribution of power) are more reluctant to delete others' links, probably because in these countries, where important decisions are made by few powerful people, editors may not feel to have the authority to delete someone else's work. Similarly, Hara, Shachaf, and Hew (2010) examined users' behavior in different discussion spaces in Wikipedia, using content analysis on 120 talk pages in different namespaces of four Wikipedias characterized by different culture and size, and specifically English (western, big), Hebrew (western, small), Japanese (eastern, big) and Malay (eastern, small). In particular, they concluded that postings expressing courtesy, such as apologies, regrets, appreciation or greetings, were more frequently found in large and Eastern Wikipedias, characterized by a high Power distance, which is associated to high politeness.

Despite the growing amount of research carried on about Wikipedia and its social practices, only few works started to analyze the formation of collective memories in

the online encyclopedia (Ferron & Massa, 2011b; Pentzold, 2009), and this dissertation is aimed at filling the gap in the study of the representation of traumatic events in Wikipedia.

2.4 Wikipedia as a global memory place

It can be argued that Wikipedia is one of those Web 2.0 services that answer the contemporary need of recording, keeping track, and archiving information (Hoskins, 2009a). In fact, many of the articles related to the most shocking contemporary traumatic events have been created just few hours, if not minutes, after the beginning of events (Ferron & Massa, 2011b; Keegan, 2011). For example, the Wikipedia article on "7 July 2005 London bombings" was initiated at 09:18, only few minutes after the first bombings. Just in the course of that day, the article was edited 2581 times. The associated talk page was created on July 7 at 09:59 and received 626 contributions in its first day.

According to Pentzold (2009), Wikipedia's processes of discussion and article construction can be seen as the discursive formation of memory, or in other terms, as the transition from communicative memory, which is interactive, informal, non-specialized, reciprocal, disorganized and unstable, to cultural memory, which is formal, well organized and objective (Assmann, 1995). In this sense, in Wikipedia memories are formed through the social interplays between users, and through their interactions with the digital medium, reconfirming remembering as a situated activity (S. Brown & Hoskins, 2010; Hirst & Manier, 2008).

From this perspective, in the case of traumatic events, editing activity can be interpreted as a sign of active participation in remembrance, where the collection of different accounts and pieces of stories can dialogue with each other and compose a unique representation. In Wikipedia, the different pieces of information added to an article or talk page, or even deleted from it, are archived and accessible in the revision history, thus serving a narrative function and telling the unique history of a page, defining what will or will not be included in the final shared representation of an event.

The dynamics of collective remembrance take place not only in the public, but also in the private sphere, for examples in family dinners, class reunions, visits with old friends, and it is not sufficient to study the past and understand its representation prescinding from these communities of memory. As Irwin-Zarecka (1994, p. 56) pointed out, "how people attend to the past, if at all, and how they make sense of it is very much grounded in their experience". Considering the novel relationship between the public and the private in the new media context, Wikipedia can be seen as laying in the middle way between the public discourse and the private interactions. Indeed, while its pages are accessible to everyone and visited by an increasingly number of people every month (Zickuhr & Rainie, 2011), the dynamics of communication taking place in the platform are in the form of everyday interactions. On the talk pages, the individual points of view and private thoughts of the participants are discussed and negotiated to reach a consensus. According to Irwin-Zarecka, if collective memory is "a socially articulated and socially maintained "reality of the past", then it also makes sense to look at the most basic and accessible means for memory articulation and maintenance – the talk" (1994, pp. 54-55). To this end,

Wikipedia is particularly suitable because the communicative interactions are registered, archived and publicly accessible through the revision history pages.

Technologies and their associated socio-technical practices play a key role in shaping how memory is formed and constructed, as noted for instance by Bowker (2005), Van House and Churchill (2008) and Garde-Hanse et al. (2009). This is true also for Wikipedia in its interpretation as a global memory place. For instance, detailed policies and guidelines provide fundamental behavioral rules that influence the way articles are written. One of the most relevant is the "neutral point of view", meaning that articles should be accurate, state verifiable information, provide authoritative references and be written proportionately and without biases ("Wikipedia:Neutral point of view", 2001). Moreover, being an online encyclopedia, Wikipedia does not promote original research, advertising, personal opinions and memorials of deceased friends, acquaintances or relatives ("Wikipedia:What Wikipedia is not", 2001). These elements indicate that the primary objective of Wikipedia is to be an online encyclopedia, and therefore its interpretation as a global memory place must acknowledge that the rules, the social practices and the technologies that are in place in the Wikipedia socio-technical system strongly influence how users write articles, even when interpreted in the context of collective memory formation. Nevertheless, Wikipedia articles and talk pages seem to accomplish also a cultural function which serves the need of the community in the present, for example providing ways to express grief and mourning, as reported in Chapter 3. Hence Wikipedia becomes an interesting playground for the study of collective memory building processes, providing unobtrusive access to the text of communicative interactions between users who work together to build an accurate and shared representation of the past, as showed in Chapters 4 and 5.

It is a place where public discourse take place, and most importantly, the written controversies are available in the talk pages associated to articles, and even when a discussion is over, they are archived and remain available to everyone. This is an extensive amount of tangled arguments, issues debates, interpretations, claims, critics, agreements, which scientists can tap into in order to study the dynamics of memory also from a longitudinal point of view, as described in Chapter 6.

Chapter 3

Commemoration of traumatic events in Wikipedia

3.1 The functional approach to collective memory and commemoration

Theoretical research emphasized the importance of the functional nature of collective memory and its close connection with the collective identity of a community (Harris et al., 2008; Hirst & Manier, 2008). Taking into account the growing interest of psychology for collective memory processes, Hirst and Manier (2008) argued that a functional perspective could help to distinguish collective memories from "shared memories", which have little to do with the collective identity of a group. Indeed, one of the major functions of collective memory is to satisfy the needs of the communities in the present such as, in the first instance, the construction and the support of groups' identity, cohesion and continuity (Harris et al., 2008).

Among the complex processes through which collective memories are formed and accomplish their functions, commemoration plays a crucial role. As a matter of fact, commemoration offers groups' members the possibility to elaborate past events through the creation and negotiation of different narratives, interpretations and perspectives (Y. Zerubavel, 1995).

According to Wang (2008), commemoration is also strictly connected to other functions of collective memory: emotional bonding, therapeutic practice and directive function. With regard to the first, Wang proposed that collective memory processes, also through commemorative ceremonies, may help to strengthen the emotional ties among group members, contributing to develop a sense of closeness, cohesiveness and solidarity in the community. Especially with regard to traumatic events, commemorative rituals, through memory retelling and externalization, allow groups to carry out coping strategies and to recover from traumatic experiences, accomplishing a therapeutic function. Communities can also reflect upon their past traumas, directing future actions by learning from their mistakes, in order to avoid the same traumatic experiences to happen again.

Generally, acts of commemoration consist of monuments, rituals and ceremonies. In recent years, the widespread diffusion of Internet technologies and social networking sites (boyd & Ellison, 2007) has led to the propagation of online memorials devoted to the preservation of collective memory and to the commemoration of traumatic events. Recent and older collective upheavals such as the September 11 attacks of 2001, the Hurricane Katrina in 2005, or the earthquake

in Abruzzi, Italy in 2009, have been widely commemorated through digital archives and social networking sites.

Arthur (2008) analyzed the online commemoration practices taking place in the September 11 Digital Archive (http://911digitalarchive.org), the Hurricane Digital Memory Bank (http://www.hurricanearchive.org), and the World is Witness geoblog (http://blogs.ushmm.org/worldiswitness). The September 11 Digital Archive and the Hurricane Digital Memory Bank are two examples of digital archives, which rapidly spread in the last years, and basically are online databases where people can upload and browse their personal thoughts and stories, images, videos or other files. Slightly differently, the World is Witness geoblog is a commemorative project with the specific purpose of increasing people's consciousness about crimes against humanity, by geographically locating the sites of interest, and thus not focusing on a specific traumatic event but on multiple episodes of violence.

Acknowledging the cathartic function pursued by these websites through the storytelling and social sharing of emotions enacted by their users, Arthur (2008) recognized that in fact they do not create collective memories, but are best to be interpreted as sites of "collected memories" (p. 8). Although these digital environments may in the long term contribute to a better understanding of past events, their individual fragments do not interweave each other to form a coherent narration, but on the contrary they form "a picture full of gaps or made up of fragments" (p. 16).

Micalizzi (2012) analyzed the content of Facebook groups dedicated to the commemoration of the earthquake in Abruzzi, Italy, in 2009, and interviewed 25 users to better understand their motivations for participating in the groups' activities. She identified four main functions in the users' practice of sharing their experiences and feelings about the earthquake. First, a pragmatic function, conveyed by posts aimed at finding news about relatives or friend living in the region, promoting solidarity actions and providing useful information about the official emergency organization plans. Second, a socio-affective function, which can be found in messages expressing anger, frustration, solidarity and empathy towards the victims. Third, an auto-therapeutic function, which can be found, for instance, in people posting messages about a similar trauma they experienced in the past. Finally, a memorative function, through which users leave their personal anecdotes, remember the victims and document the tragedy in order to preserve its memory for the future. In line with Arthur's (2008) conclusions, Micalizzi argued that Facebook groups' activities cannot be considered in the framework of collective memory building processes, because users' participation is fragmentary and plural, and lacks two important characteristics of collective memory: order and selection. In the Facebook groups dedicated to the earthquake there is no dialogue between single posts, which are rather autonomous and independent. On the contrary, collective memory practices are the result of complex and dynamic processes of selection and discussion of relevant information, which are ordered into coherent narratives shared by a community, as it happens in Wikipedia articles and talk pages (Pentzold, 2009).

Similarly to Arthur (2008), Recumber (2012) analyzed the September 11 Digital Archive and the Hurricane Digital Memory Bank, interpreting them as sites of prosumption, where users can both consume and produce content. The author studied the messages uploaded by users to these two digital archives in order to better understand online commemoration practices, and found that the stories recounted are often related to emotions and healing, and therefore have a therapeutic

significance. Acknowledging the healing power of memorials, such as the Vietnam Veterans Memorial, Recumber focused specifically on the therapeutic function of digital archives and memory banks. These websites flourished to allow users to reflect upon a traumatic experience, and to empathize with victims and survivors through the spontaneous sharing of their personal stories and thoughts. Such online platforms, through the interaction and active participation of their users, produce a wide network of entangled threads, posts, photos which seems fragmentary, built on the model of a database rather than a narration.

Recumber recognized different functions of online commemoration practices in the analyzed digital archives. One of these functions is commemorating the dead, usually particular victims such as a brother, a friend, a mother. In this case users describe how they experienced the loss or remember significant aspects of the victims life and personality. Another function accomplished by the digital archives is a therapeutic one, through which users, telling their personal stories, refer to the psychological consequences they went through, such as depression, distress or sleep disturbances, seeking to overcome the traumatic experience through the social sharing of their feelings. Connected to this therapeutic function, the author found a substantial amount of messages with a spiritual or religious content. On the other hand, the commemoration activities allowed users to vent their frustration, fear and sadness, directing their anger toward the state and the local government in the case of Hurricane Katrina, and toward the terrorists in the case of the September 11 attacks. However, in spite of the therapeutic function accomplished by digital archives, the author observed that the recollection of the past through collective memory practices needs to be an active and constructive process, and not only a fragmentary collection of pieces of information. Indeed, memory banks appear to be exploited more as individual healing tools where prosumers leave their messages and browse others', in order to work through the traumatic experience. Hence, digital archives appear to have missed the opportunity to foster a "collective experience" (p. 547), reflecting an individualistic and therapeutic attitude, with users contributing in an inward and atomized manner. On the contrary, as already discussed in Chapter 2, the online encyclopedia Wikipedia can be interpreted as a global memory place where collective memory building takes place, because the individual pieces of information are selected and ordered into a coherent narration, through users' discussion and active participation in remembrance.

The digital memory practices considered so far can all be interpreted as commemoration activities, aimed at the social sharing of personal feelings and experiences in order to not forget the past. Nevertheless, when it comes to commemorating there is a widespread need to do so on specific days and at specific intervals, and this applies to private as well as to public events, traumatic or not (Frijda, 1997). During the commemoration of traumatic events, which are usually held at regular intervals starting from a year after the original episodes, people reexperience the trauma and devote their attention to remember it. Similarly to other recurring events, such as birthdays, funerals or Christmas, commemorations are occasions to think about some persons or historical happenings in specific and publicly recognized days which bear a temporal relation with the events being commemorated. According to Frijda (1997), in this way people can satisfy their need for a temporal orientation, structuring time by identifying precise demarcation points and thus giving an order to its flow.

The re-appropriation of the past taking place during commemorations is part of the definition of the collective identity of a community, which is deeply related to its collective memory. During commemorations people take the opportunity to understand past events and negotiate their meaning. Thus, commemorations reinforce the need for temporal orientation, for the appropriation of the past, and for the definition and maintenance of a group's identity and cohesion.

Frijda recognized three major features defining the functions of commemoration. The first is related to the need for temporal orientation, according to which commemorative ceremonies usually happen during publicly recognized days which are socially shared by a community. The second feature of commemoration is its symbolic and communal nature. Commemorating activities are usually shared with others: people get together in front of a memorial to listen to a public speech or to a band playing a national hymn, and participate in other public activities. Third, the communal context in which commemorations take place relates them to rituals. In general, rituals are performed to provide order, stability and coherence, and are usually defined by a community or by tradition to pursue an emotional and a moral goal through the re-evocation of the past and the reinforcement of a group's identity.

Commemorations are also frequently emotional occasions: the emotions aroused may be most often of sorrow and grief, anger or sadness, and sometimes also of pride for a former victory or national achievement. During the ritual of commemoration, people are socially accepted by others as emotionally affected individuals, and the social sharing of these feelings allows to empathize with each other strengthening group solidarity, and to come to terms with the traumatic experience, stating its emotional and social meaning.

Finally, similarly to Wang (2008), Frijda proposed that commemoration rituals are not only motivated by the need to structure the past and to come to terms with it. Also, they usually relate the past event to its implications in the present and the future, to awake people's awareness in order to prevent the same traumatic experience to happen again.

Considering the functions of commemoration previously mentioned and interpreting Wikipedia as a global memory place, this chapter aims at introducing quantitative methods in the study of collective memory building. If Wikipedia is a place where collective memory is formed through the active collaboration of its users (Pentzold, 2009), then the typical functions of collective memory processes should emerge from the patterns of edit activity and from editors' comments in the discussion pages related to traumatic events. Focusing on these articles and talk pages, related to events such as the September 11 attacks or the 7 July 2005 London bombings, we expect to find significant boosts of users activity during anniversaries. The increased participation in the construction of a common representation of the past at regular intervals would indeed confirm the presence of commemoration activities in Wikipedia, making visible users' need to engage in the periodical re-appropriation of the past, taking part in the formation of the collective memory of traumatic events. A further analysis of meaningful comments in the discussion pages during anniversaries of traumatic events will finally highlight the presence of the characteristic functions of collective memory formation in Wikipedia.

3.2 Method

The Wikimedia Foundation maintaining Wikipedia servers provides datasets containing every edit made to every page for the different languages of Wikipedia[1]. Specifically, we developed an open source script to process an XML file including the full revision history of all pages of the English Wikipedia at 16 September 2010, and consisting of more than 400 million edits to more than 3 million articles.

An initial data exploration is presented in *Figure 3.1*, which shows a visual representation of the monthly number of edits over time for the articles on "September 11 attacks", "7 July 2005 London bombings" and "Chernobyl disaster". It is possible to visually observe the presence of spikes in users' edit activity around anniversaries. For example, the article on "September 11 attacks" received a very large number of edits during the months of September 2005 and 2006. In the article "7 July 2005 London bombings", edit activity clearly increased in July 2006, 2008 and 2009, while for the article on "Chernobyl disaster", occurred in the night between 25 to 26 April 1986, a pronounced peak of activity can be observed by the 20th anniversary in 2006. The graphs show also few increases in edit activity occurring outside anniversary periods that may be due to different reasons, such as news updates about the events triggered by new information about trials, or emerging debates among Wikipedia users.

---

[1] Wikipedia datasets are available at http://dumps.wikimedia.org. The datasets on which the logistic regression analysis has been run, along with the scripts employed to extract edit activity and released as open source, are available at http://sonetlab.fbk.eu/data.

Figure 3.1 Number of edits per month over time to the articles on September 11 attacks, 7 July 2005 London bombings, and Chernobyl disaster.

While the graphs reported in *Figure 3.1* are visually revealing and hint at the presence of regularities in edit patterns, the purpose of this chapter is to test if these patterns are statistically significant. In other words, the aim is to quantitatively verify if spikes of activity are a characteristic of pages related to traumatic events that allow to discriminate between these and other Wikipedia pages.

For every article and, separately, every talk page, we extracted from the XML file the total number of edits occurred since their creation on Wikipedia (column "Total edits" in *Table 3.1*), maintaining also the information about the number of edits occurred each day, in order to group edits made around anniversaries as explained in

the following. We also computed the average number of edits per day for each page (see *Table 3.1*).

Taking into account the unfolding of collective memories processes, only pages with a substantial amount of participation by different users were considered. For this reason, we computed the number of *unique editors* for every page and excluded from our dataset articles and talk pages edited by less than 50 different users. This filtering reduced the dataset to a representative core of 396,412 articles and talk pages. With a separate process, we looked for traumatic events represented in Wikipedia as specific article pages. In order to find traumatic events, we exploited the large structure of Wikipedia internal categories such as "Events by topic", "Man-made disasters", "Earthquakes", and lists of important historical events available on other online encyclopedias and digital databases such as Britannica Online Encyclopedia (2011), History Central (2011), and Information Britain (2011). In this way we were able to identify 88 pages (57 articles and 31 talk pages) related to traumatic events edited by at least 50 different users. The remaining group of pages hence represented other Wikipedia pages, i.e. pages that are not related to traumatic events but can be about anything else, such as for example "William Shakespeare", "Infinitesimal calculus" or "Italy". This comparison group consisted of 377,651 articles and 18,673 talk pages. The relative difference in the number of articles and talk pages is due to the fact that the talk page related to an article is created only if at least one Wikipedia user intends to discuss the article's content. Moreover, talk pages tend to have less contributors than articles and hence requiring at least 50 different editors filtered out many more talk pages than articles. The complete list of articles and talk pages, along with the computed measures, is available in the released dataset.

Since the focus was on the construction of collective memory through people collaboration, we excluded from the full dataset edits made by *bots*, automated tools that carry out repetitive tasks to maintain Wikipedia pages. At September 2010 Wikipedia counted more than one thousand bots ("Wikipedia:Bots", 2007), commonly used to perform minor edits, such as fixing unsigned comments, correcting broken links, reporting possible copyright violations, etc. The Wikipedia articles and talk pages included in the original dataset received in total 211,364,715 edits, 7.77 % of which were made by bots.

For every article and talk page we also collected the date of first edit, i.e. of creation of the article or talk page. Wikipedia pages about recent traumatic events tend to get created the same day or few days after the event (Ferron & Massa, 2011b; Keegan, 2011). As previously reported in Chapter 2, the Wikipedia article on "7 July 2005 London bombings" was initiated at 09:18, only few minutes after the first bombings, and received 2581 edits during the first day. Our dataset contained 30 articles about traumatic events happened after Wikipedia was launched in 2001. Out of them, 23 (77%), were created within two days after the event (18 were created the same day). Similarly, out of 21 talk pages about traumatic events that happened after Wikipedia was launched in 2001, 13 (62%) were created within two days after the event (11 were created the same day).

As introduced above, this chapter focuses on commemoration practices of traumatic events emerging from edit activity around anniversaries in the English Wikipedia. Precisely, we considered the period of 21 days around the anniversary of the event (a time window of ten days before and after the anniversary), to take into account users activity in preparation for the anniversary day and the subsequent activity in the immediately following period. For example, for the second anniversary

of the London bombings occurred on 7 July 2005, the edits made to the related Wikipedia pages since 27 June 2007 up to 17 July 2007 were considered.

The goal of this chapter is to show periodical boosts of activity during anniversaries in pages about traumatic events comparing them to other Wikipedia pages. Specifically, for the latter ones, which do not necessarily refer to events with specific dates, we considered the period of 21 days around the anniversary of the creation date as a baseline over which performing the comparison. This assumption is reasonable since, as explained above, pages about recent traumatic events occurred after the launch of Wikipedia in 2001 tend to be created soon after the events (Ferron & Massa, 2011b; Keegan, 2011). In this case, the page creation date and the event date are very close if not the same. Considering the edit activity during the anniversary of the creation date for the comparison group allowed to enlarge the dataset and base the results on a larger evidence.

Focusing on edit patterns around anniversaries, we filtered out the edits made during the first six months after the creation of a page, a period in which the discussion may be very heated but cannot be still considered related to commemoration activities.

In summary, the dataset consisted of 57 articles and 31 talk pages about traumatic events and other 377,651 articles and 18,673 talk pages, all of them edited by at least 50 different users.

In order to compare Wikipedia pages, for each of them two measures were computed: the average number of edits per day excluding, as already mentioned, the first six months since page creation, and the average number of edits per day during the anniversary periods (see *Table 3.1*).

Pages related to traumatic events	Unique editors	Total edits	Average edits per day	Anniversary edits	Average anniversary edits per day	Average anniversary edits per day / average edits per day
Article						
September 11 attacks	4697	14813	4.62	1969	10.70	2.32
Attack on Pearl Harbor	3223	8299	2.59	1011	5.35	2.07
John F. Kennedy assassination	3222	8057	3.11	794	5.40	1.73
Chernobyl disaster	3795	7740	2.65	1164	6.93	2.62
Hurricane Katrina	5969	6826	3.52	579	5.51	1.57
Talk page						
September 11 attacks	2204	17468	7.22	2198	15.48	2.14
Attack on Pearl Harbor	424	1689	0.57	120	0.71	1.26
Mean and Standard deviation	**740.53 (sd = 1108.50)**	**1531.16 (sd = 2872.90)**	**0.63 (sd = 1.10)**	**182.21 (sd = 365.09)**	**1.36 (sd = 2.43)**	**2.20 (sd = 1.35)**

Other Wikipedia pages	Unique editors	Total edits	Average edits per day	Anniversary edits	Average anniversary edits per day	Average anniversary edits per day / average edits per day
Article						
Gregor Schlierenzauer	125	242	0.18	13	0.20	1.17
Carlo Gambino	325	691	0.24	40	0.27	1.12
Master P	1227	2432	0.87	138	0.94	1.08
Christine Keeler	188	269	0.10	14	0.19	0.99
Quentin Tarantino	2759	4878	1.44	260	1.38	0.95
Talk page						
Sassanid Empire	128	394	0.20	19	0.18	0.90
Zionism	431	3239	1.22	170	1.16	0.94
Mean and Standard deviation	**180.59 (sd = 296.56)**	**359.66 (sd = 720.98)**	**0.17 (sd = 0.30)**	**21.36 (sd = 49.53)**	**0.20 (sd = 0.50)**	**1.13 (sd = 0.95)**

Table 3.1 Few examples of the information computed for every page of the dataset.

As an example, the first row of *Table 3.1*, referring to the article about the September 11 attacks, shows that excluding the first six months, at 16 September 2010 the article counted 14,813 edits by 4697 different editors. During the 21 days periods around the nine anniversaries occurred since the page creation in 2001, the page was edited 1969 times. Hence, the average number of edits per day was 14,813/3207 = 4.619, while the average number of edits per day during anniversaries was 1,969/184 = 10.701. Therefore, while on average the article was edited more

than four times per day, during anniversary periods Wikipedia users modified it more than ten times per day.

For comparative purposes, for every page we computed an additional variable, the ratio between the *average number of edits per day during anniversaries* and the *average number of edits per day*. Intuitively, if there was no difference between edit activity during anniversary days and other days (i.e., if the number of edits was almost the same every day), this ratio would be around 1. If it was higher than 1 (for example, 3) for a certain page, this would suggest that an edit during the anniversaries is 3 times more probable than in any random day. For example, *Table 3.1* shows that the ratio *average anniversary edits per day / average edits per day* for the article on September 11 attacks was 2.32, suggesting that an edit was 2.32 times more likely to occur during one of the nine anniversaries than in any random day. Over articles and talk pages related to traumatic events, the mean ratio was 2.20 (sd = 1.35), while for other Wikipedia articles and talk pages the mean ratio was 1.13 (sd = 0.95). This suggests that edits to pages related to traumatic events are 2.20 times more likely to occur during anniversary periods, while other pages are edited more constantly along the years.

In order to verify if this ratio could significantly distinguish between pages related to traumatic events and other pages, we applied a logistic regression model with *traumatic* as binary dependent variable and the ratio *average anniversary edits per day / average edits per day* as independent variable. The basic idea of this approach is to predict if a page is related to a traumatic event or not depending on the ratio of edits occurred during the anniversaries and in other days.

3.3 Results

Specifically, we ran two separate logistic regression analyses for articles and talk pages. With regard to articles, the regression coefficient for the ratio was statistically significant (0.315; p<.001). This means that an increase of one unit in this variable increases the log-odds in favor of an article being related to a traumatic event by an estimated 0.315 with a confidence interval of less than 0.1%. As for talk pages, the regression coefficient for the ratio was also statistically significant (0.223; p<.001). In other words, the values assumed by the ratio *average anniversary edits per day / average edits per day* can significantly discriminate between pages related to traumatic events and others, meaning that pages characterized by a high relative amount of edits during anniversaries are more likely related to traumatic events.

Although the focus of this chapter is on commemoration of traumatic events during anniversaries, a comparative analysis on non traumatic events was also performed. Precisely, we looked for events of the twentieth century of manifest historical importance which were inherently non traumatic and not related to traumatic events. We did this by following the same procedure adopted for traumatic events, that is by exploiting Wikipedia internal categories such as, in this case, "Cultural events", "Sports events", "Festivals", "Ceremonies" and more and also lists of important historical events available on Britannica Online Encyclopedia (2011), History Central (2011), and Information Britain (2011). The filter on the minimum number of unique editors previously used was required for these pages as well, and only 9 articles and 4 talk pages were found to have at least 50 different contributors (see *Table 3.2*). On the other hand, in the previous analysis 57 articles and 31 talk

pages passed the filter. While the comparison cannot be considered definitive, this difference seems to suggests that non traumatic events receive less attention in Wikipedia when compared to traumatic events. However, few non traumatic events such as "Woodstock Festival" or "Apollo 11" attracted a large number of editors, respectively 2274 and 2244 (see *Table 3.2*), while traumatic events of comparable historical importance, represented within Wikipedia in articles such as "Chernobyl disaster", "Attack on Pearl Harbor", or "John F. Kennedy assassination" received contributions respectively by 3795, 3223, and 3222 different editors.

Pages related to non traumatic events	Unique editors	Total edits	Average edits per day	Anniversary edits	Average anniversary edits per day	Average anniversary edits per day / average edits per day
Article						
Apollo 11	2244	4803	1.47	534	2.83	1.92
Sputnik 1	859	1495	0.46	303	1.57	3.41
Sputnik 2	186	279	0.10	33	0.20	2.05
Explorer 1	211	511	0.17	107	0.64	3.71
Treaties of Rome	205	357	0.12	52	0.31	2.54
Treaty of Lisbon	1287	2879	2.46	103	2.45	0.10
Convention on Biological Diversity	134	229	0.07	16	0.08	1.22
Treaty of Nice	160	232	0.08	9	0.05	0.68
Woodstock Festival	2274	4560	1.42	562	2.97	2.10
Talk page						
Apollo 11	227	509	0.16	103	0.54	3.33
Sputnik 1	67	140	0.05	25	0.14	2.98
Treaty of Lisbon	318	1319	1.13	77	1.83	1.62
Woodstock Festival	139	278	0.09	54	0.28	3.30
Mean and Standard deviation	**639.31** (sd = 797.33)	**1353.15** (sd = 1664.61)	**0.60** (sd = 0.77)	**152.15** (sd = 191.08)	**1.07** (sd = 1.11)	**2.30** (sd = 1.00)

Table 3.2 List of pages related to non traumatic events of historical importance with at least 50 unique editors.

Table 3.2 presents the data already reported for traumatic events (*Table 3.1*) considering non traumatic events. Since the mean ratio *average anniversary edits per day / average edits per day* for these pages is 2.30, these articles and talk pages related to non traumatic events seem to follow similar commemoration patterns during anniversaries when compared to trauma related pages. However, some of these non traumatic events were also of crucial importance for the history of mankind. The articles and talk pages of "Sputnik 1", "Sputnik 2", "Explorer 1" and "Apollo 11" all refer to one of the greatest technological races in human history, the Space Race, which is closely related to the Cold War. Indeed, Sputnik 1 and Sputnik 2 were respectively the first and the second artificial satellites launched by the Soviet Union in 1957, followed by Explorer 1, launched by the United States in 1958. Ten years

later, in 1969, United States launched Apollo 11, the spaceflight which landed for the first time two men on the Moon and which can also be considered one of the most significant events of the Space Race between United States and the Soviet Union. Moreover, if on the one hand these singular events cannot be considered as traumatic, they are still indirectly connected to a fundamental period of unprecedented tension between two superpowers, USSR and USA, over nuclear arms. Thus, it is not surprising that the associated pages are mostly edited during anniversaries, since these events are pivotal for the human history.

However, the small number of pages about non traumatic events edited by at least 50 editors suggests that, in general, non traumatizing events may receive less attention in Wikipedia, except for the ones of manifested historical importance reported above. Moreover, the exiguous number does not allow to test for significant differences between the edit activity around anniversaries of traumatic and non traumatic events.

The purpose of this chapter was to show that the investigation of collective memory formation of large numbers of traumatic events in Wikipedia is possible and interesting. In other words, the analysis performed aimed at opening the way to the study of collective memory of traumatic events in Wikipedia, showing that commemoration patterns emerge from users' edit activity during anniversaries. Indeed, the ratio *average anniversary edits per day / average edits per day* could significantly discriminate between pages related to traumatic events and other pages, meaning that those articles and talk pages characterized by a high relative amount of edits during anniversaries are more likely related to traumatic events.

Commemoration is one of the most salient characteristics of collective memory processes, especially in the case of traumatic events, in which it can perform several functions to meet the needs of a community in the present (Frijda, 1997; Wang, 2008). In the following section these functions are discussed, taking into consideration the most meaningful comments left by Wikipedia users on the talk pages during anniversaries of traumatic events.

3.4 Discussion

Considering Wikipedia as a place where the negotiation of different, and sometimes contrasting, interpretations of the past takes place, the analysis performed in this chapter is a first step towards the quantitative study of online collective memory processes on a large scale. Specifically, we considered the number of edits in time to Wikipedia pages to verify if the relative amount of edit activity during anniversaries can significantly distinguish pages related to traumatic events from other pages.

In order to verify if the observed pattern is a characteristic capable of discriminating pages as related to traumatic events or not, two logistic regression models were applied to articles and talk pages. Basically, these models allowed to predict if an article or talk page was related to a traumatic event depending on the value assumed by the independent variable.

Results showed that the relative amount of edit activity during anniversaries is a statistically significant predictor in discriminating pages related to traumatic events from other pages. In other words, pages with a high relative amount of edits occurred during anniversaries are more likely associated to traumatic events. These results suggest the presence of increased discussions and other activities related to

commemoration carried out by Wikipedia users on pages about traumatizing events, during their anniversaries.

While the overall number of contributions to Wikipedia articles about traumatic events may tend to decrease over time as the representation of the facts become more stable and crystallized, the regular peaks of activity during anniversaries, even several years after the originating trauma, highlights the "lack of closure" of such events (Sturken, 1997, p. 63). In Wikipedia, the creation of a shared narrative framework is continuously ongoing, as new information is recovered and interpreted. But anniversaries in Wikipedia seem to become an occasion for further re-interpretation of the past and the re-construction of a shared memory, as well as a moment for grieving and mourning (E. Zerubavel, 2003). Remembering is a dynamic process that requires order and selection (Jedlowski, 2001; Sturken, 1997), and building a shared narration means to make decisions about what is worth recovering. Often, these decisions are highly controversial and may result in conflicting opinions, heated debates, and sometimes edit wars.

The fervid activity that can be noticed on many articles and talk pages about traumatic events in Wikipedia shows that people are eager to record an accurate representation of these events. However, on the talk pages these contributions seem to go beyond the encyclopedic intent of Wikipedia, and recall several functions highlighted by scholars in the past literature (for example, Arthur, 2008; Frijda, 1997; Micalizzi, 2012; Recumber, 2012; Wang, 2008). The following comments are part of different discussions extracted from the talk page of the "September 11 attacks" article during the fifth anniversary, in September 2006, and show particularly the socio-affective and memorative functions of collective memory:

> **Memorial**
> Is there any plans that any of the Wikipedia Adiministrators would make a memorial page for the events of that tragic day? Also will this be able to be a permanent memorial online? --82.47.145.146 21:39, 4 September 2006 (UTC)
> There already is one, http://sep11.wikipedia.org but it's in constant danger of being removed. --Golbez 19:44, 7 September 2006 (UTC)
>
> I find it extremely upsetting and unfortunate that there were so many unneccessary deaths involved in the September 11 "attacks". […]
> I cannot convey enough how tragic it really is to know there are so many lives that are lost today because of the American goverment's attempt to control its citizens through fear and prejudice against another society. To me, it is far worse for the families of victims of 9/11 than for families of soldiers (although I still sympathize and feel for the families, but hear me out) because in war, we all understand that there are casualties. We understand, however difficult that may be, that there will be many circumstances in which lives are lost no matter how prepared and well-trained the individuals may be.
> But what makes 9/11 far worse beyond any comprehension I have been able to gather, is that the people that were killed in the WTC on September 11, 2001 did not die in a time of war. They did not die knowing that they were in a dangerous situation, serving their country against a powerful threat to their existence and freedom. Yes, there were many rescue personnel on site (nearly 400 fire fighters lost their lives), but they also were not sent into a war zone at the time. They were sent in because they thought there would be a way to rescue those trapped in the buildings. They were not given the forewarning that the buildings were planned to collapse.

[…] these are just the thoughts of one extremely infuriated individual. Take a second and think about it: Does it not make you angry and bewildered beyond belief that nearly 400 rescue workers died saving the lives of others while members of municipal government were safely tucked away with prior knowledge of the collapse?

I did not know any of the people whose lives were lost in 9/11 yet I feel the urge to convey just how hard this hits me still. I'm not talking about the attacks. I'm not talking about the sudden feeling of lost national security. That is not what scares and angers me. These lives were viewed as just a means to an end for the government. They did not think twice about sending in people with families, children, mothers, fathers, cousins, friends, etc. They saw then as a tragic loss caused by malicious terrorists who the American government will make sure to seek revenge on. […]

My heart goes out to the families, friends and loved ones of anyone who lost their life in the 9/11 attacks. […]

Please, do your part in uncovering the truth as I have begun to do. […] --207.210.23.114 16:00, 10 September 2006

5 Year Anniversary
Tonight in Australia is the 5 Year anniversary of the Sep 11 attacks. I lost my mum a few months before Sep 11 to cancer, and I know what grief is like. My prayers are with those who are related/friends with the dead of Sept 11. May the terrorists be defeated and brought to justice! May God bring about His justice on Earth! Geelin 13:34, 11 September 2006 (UTC)

Anything for the families?
[…] I see no reference in the article to the effects on the families of those who have died, etc.
[…] I'm suggesting that a section about the effects on families and the grief perhaps be added. […] KBecks 19:16, 10 September 2006 (UTC)

Bereavement is universal; we don't have a section for that on every war and disaster since what people feel is pretty much the same in every case, and we already have a general article for that. […] --cloviz 01:08, 11 September 2006 (UTC)
Maybe we can add something about the 36 children whose fathers died that day, even as they (the kids) were still in their mothers' wombs. Cerebral Warrior 11:09, 11 September 2006 (UTC)

Terrorist attacks
I am going to make it plain and simple. Do not add "terrorist" to the lead back. It will only be removed. --Cat out 09:40, 11 September 2006 (UTC)

There is not better description of the action. Their actions were ones of terrorism.--MONGO 09:48, 11 September 2006 (UTC)

[…] The term, "terrorism" is appropriate in this case. […] There is wide agreement among countless reliable sources that 9/11 was an act of terrorism. This has been discussed before, with consensus to leave the word terrorism in the intro. --Aude (talk contribs as tagcloud) 10:02, 11 September 2006 (UTC)

[…] I don't understand how the murder of 3000 innocent people can NOT be called "terrorist." Please, today is 9/11. Let us pray for the souls of the deceased instead of insulting their memory by not terming those who so cruelly killed thousands of fathers, mothers, brothers, sisters, friends, as terrorists. Cerebral Warrior 11:08, 11 September 2006 (UTC)

[…] Look at it this way- if today was Holocaust Memorial Day, would you guys be insisting that a Holocaust denial theory proposed by some Moslem dictator be included in the article about the Holocaust? No, you wouldn't because that would be an insult to the 7 million people who died in Nazi gas chambers. Well, today is a day when we remember the 3000 innocent men, women and children whose lifes were snuffed out by a new breed of fascists-Islamofascists. Referring to their murders as anything other than "terrorists" would be an insult to the deceased, who do not deserve to be mocked simply because they were Americans. If you are an anti-American, that's your wish, but please don't insult the departed. Cerebral Warrior 13:43, 11 September 2006 (UTC)

[…] my sympathy and prayers to those who mourn this day.--Doc 14:08, 11 September 2006 (UTC)

[…] I am saddened by the fact that Wikipedia is not doing enough to mourn this day- including not even mentioning the fifth anniversary on the mainpage. Cerebral Warrior 14:36, 11 September 2006 (UTC)

[…] Spare a thought for those whose lives were torn apart that day. Cerebral Warrior 14:39, 11 September 2006 (UTC)

[…] Please don't inflame the situation. This is a sensitive topic, especially on the fifth anniversary of the event. […] --Guinnog 14:55, 11 September 2006 (UTC)

The comments reported above[2] were made on the talk page of the article about the September 11 attacks during the days surrounding its fifth anniversary, and are the expression of different functions of collective memory practices. First, they convey users' feelings of anger, frustration, fear and lost of security, as it happened in the September 11 Digital Archive and the Hurricane Digital Memory Bank analyzed by Recumber (2012). Similarly to these digital archives and to the Facebook groups dedicated to the earthquake in Abruzzi studied by Micalizzi (2012), in the Wikipedia talk page about the September 11 attacks some users prayed for the victims and their families, remembered a dead relative, asked for justice and for respect during the anniversary day, thus accomplishing a memorative function. The verbalization and externalization of personal feelings and experiences is typical of commemoration and it has been found to have a therapeutic effect, helping people to get over the trauma and to cope with it (Arthur, 2008; Micalizzi, 2012; Recumber, 2012; Wang, 2008). To this regard it is important to note that, according to Wikipedia guidelines, the purpose of Wikipedia is to build an encyclopedia, and for this reason talk pages should be used to discuss improvements to the related article page. Nevertheless, on

---

[2] The full text of the comments about "Memorial" and "5 Year Anniversary" can be found on the talk page revision history, respectively at
http://en.wikipedia.org/w/index.php?title=Talk:September_11_attacks&diff=prev&oldid=74919470
and
http://en.wikipedia.org/w/index.php?title=Talk:September_11_attacks&diff=next&oldid=75092050.
The complete discussions about "Anything for the families?" and "Terrorist attacks" are archived on Wikipedia at
http://en.wikipedia.org/wiki/Talk:September_11_attacks/Archive_21#Anything_for_the_families.3 F and
http://en.wikipedia.org/wiki/Talk:September_11_attacks/Archive_21#Terrorist_attacks.

pages about traumatic events, it is possible to find signs of emotional participation related to commemoration occurring around the anniversaries.

Second, the heated discussion (available in its entirety in the talk pages' archives and revision history) which took place during the fifth anniversary about the possibility to add the word "Terrorist" to the title of the main article refers to the periodical re-appropriation of the past highlighted by different authors (for example, Frijda, 1997; Irwin-Zarecka, 1994; Y. Zerubavel, 1995). Indeed, commemoration offer the occasion not only to reaffirm the collective memory of an event, but also to negotiate highly different opinions about it and new interpretations of the past.

Another motivation for engaging in sensemaking activities and creating public narratives about traumatic events is connected to the directive function of memory analyzed by Frijda (1997) and Wang (2008). The following comments are part of a discussion which took place on the talk page of the Virginia Tech massacre article during its first anniversary, related to the possibility to nominate it as a candidate to be a "Featured article"[3]. Featured articles are considered to be the best articles in Wikipedia, following strict criteria of accuracy, neutrality, completeness and style ("Wikipedia:Featured articles", 2001). The discussion shows that users are aware of the directive function of their memory building activities, arguing that one motivation for writing a good article about the shooting is to "educate people", in order to "prevent" tragedies like this to happen again:

> **Featured Article**
> Has anyone else brought up the idea that MAYBE it might be a bad idea to list this as a featured article? I wonder how many other mentally disturbed people like that man might think "Wow, I might even get a featured article out of doing something like this!" […] The uber massive media saturation of Columbine potentially caused copycats […]. Cs302b (talk) 03:33, 16 April 2008 (UTC)
>
> So we should be quiet and pretend it never happened. Just like slavery, genocide and the Holocaust. All we're doing is making people want to be on Wikipedia and be on the news so we should burn all the books that mention these bad things so they never get these ideas. 128.227.104.129 (talk) 03:40, 16 April 2008 (UTC)
>
> Slavery, Genocide, and the Halocaust all weren't done by one single person who could do that horror on a day's notice. Report on shootings, yes. Have an article dedicated to the event, yes. […] Cs302b (talk) 03:53, 16 April 2008 (UTC)
>
> As the WP:FAC and WP:TFA nominator […] I can assure you that the intention of placing the article on the front page was to serve the public good. I believe the only way we can prevent another tragedy like this is by educating ourselves about what went wrong. This article does an outstanding job of doing exactly that. This article describes how a mom intervened when she saw troubling signs in her son and may have prevented a similar situation. I have no idea if she consulted this article or not, but I like to think that if she had she might have found some useful information. Ronnotel (talk) 04:06, 16 April 2008 (UTC)

---

[3] The complete discussion about the possibility to nominate the article about the Virginia Tech massacre is archived on Wikipedia at
http://en.wikipedia.org/wiki/Talk:Virginia_Tech_massacre/Archive_14#Featured_Article.

[…] How does this article REALLY prevent another school shooting? Featuring it is an example of what this kid was going for. […] He wanted EXPOSURE to instill FEAR. Featuring this article at best only further adds to that culture of fear and at worst reminds one lone madman that a shooting that was several months ago won't be forgotten by our media.. He'll get annual memorial services reminding people of what he did and wikipedia pages (maybe even featured)... and FAME. […] Cs302b (talk) 04:32, 16 April 2008 (UTC)

Since when is it Wikipedia's responsibility to prevent shootings? 128.227.104.129 (talk) 04:36, 16 April 2008 (UTC)

[…] I can think of no greater pursuit on Wikipedia than to create the most accurate, informative, and unbiased articles possible on topics such as these. As Rononotel pointed out above, education is the most important weapon in combating the depravity that too often beleaguers us. There's no stretch to Raul's point; the real stretch is actually believing that by eschewing the issue we can be more secure in our lives. That route entails complacency, and to argue that we shouldn't make this article prominent because it's giving him "what [he] was going for" is to completely oversimplify both his motivation and the circumstances of the tragedy. I have nothing but gratitude for those that worked to make this a fine demonstration of informativeness. Nufy8 (talk) 04:52, 16 April 2008 (UTC)

It is a good article. […] I just think that yearly massively public remembrances of something that one single person can pull off in a couple hours time should be reconsidered as there is no doubt (evidenced by his video) that he wanted to have massive public exposure.Cs302b (talk) 05:00, 16 April 2008 (UTC)

Extending this logic of complacent or passive inducement, we shouldn't feature articles about countries because it might make cause exoduses, foods because it would cause obesity, or actors because it would make people watch movies. Wikipedia articles — like movies, music, games, and other alleged depraved scourges of modern culture — don't cause people to commit crimes. Encyclopedias educate people. Educated people resist ignorance and hate. Ignorance and hate beget injustice and violence. Madcoverboy (talk) 05:06, 16 April 2008 (UTC)

Obesity, chronic movie watching (is that a major issue in the U.S.?), and exoduses are all completely off topic.. And Movies, Music, Games and those other depraved scourges of modern culture are only barely closer to the topic. I'm talking about mass media's commemorative yearly remembrances of something that ONE SINGLE PERSON can do in a short period of time with little effort.Cs302b (talk) 05:34, 16 April 2008 (UTC)

In Wikipedia pages about traumatic events, the commemoration patterns observed in users' edit activity and the emergence of collective memory functions in the talk pages during anniversaries are only two characteristic allowing to interpret the online encyclopedia as a global memory place. The third is the presence of a collective and collaborative dimension in articles' editing processes. Contrary to the individual character of many digital archives expressly devoted to commemoration (Arthur, 2008; Micalizzi, 2012; Recumber, 2012), Wikipedia is not a fragmentary collection of disconnected pieces of information. Different stories, sources, interpretations and points of view are selected and organized into coherent

narratives. The dynamic process of selection of significant elements and order into a meaningful representation is indeed essential to collective memory processes (Jedlowski, 2001; Micalizzi, 2012; Pentzold, 2009).

This chapter showed that Wikipedia pages about traumatic events receive more contributions during anniversaries. The examples of discussions reported above, occurred on talk pages of traumatic events during their anniversaries highlight the emergence of the typical functions of collective memory in Wikipedia, suggesting that collective memories processes indeed take place in the online encyclopedia. However, there may be also other reasons behind these spikes of activity. For example, it seems likely that the increased media coverage of traumatic events around anniversaries can bring Wikipedia users attention to the corresponding pages. Mass media surely play a key role in driving users' attention to specific arguments even outside anniversaries of traumatic events. For instance, a noticeable increase of edit activity (with 433 contributions) can be observed for the article about the Chernobyl disaster in March 2011, and specifically during the days in which Japan was shocked by an earthquake and a tsunami with disastrous consequences, among which the incident at the nuclear plant in Fukushima Daiichi. *Figure 3.2,* which is an updated version of *Figure 3.1*, shows the edit activity to the article "Chernobyl Disaster" until May 2011, and allows to appreciate a large spike in March 2011, which seems to be caused by the related event in Japan, which received a massive media coverage.
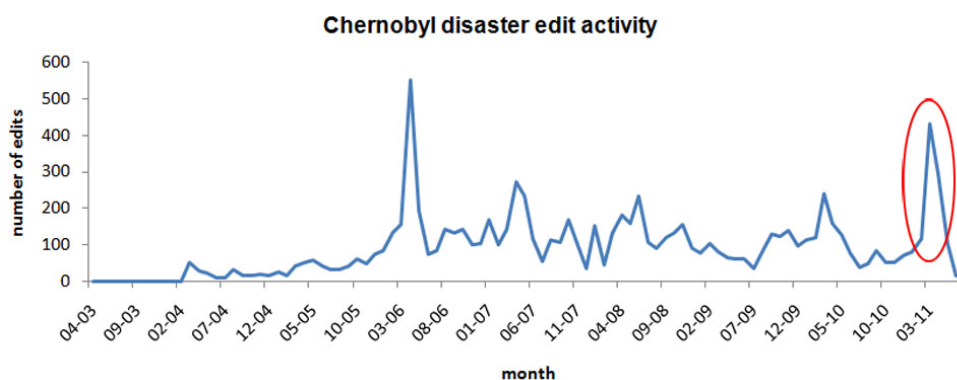

Figure 3.2 Number of edits per month over time to the article on the Chernobyl disaster.

Summarizing, the quantitative analysis presented in this chapter, together with the analysis of the content of users' comments in the talk pages during anniversaries, supports the hypothesis of the presence of commemoration practices in Wikipedia pages about traumatic events. As in the case of the article related to the Chernobyl disaster reported above, some of the peaks in the edit activity may be also due to the increased attention drew to certain events by the mass media. However, to this regard it is essential to observe that mass media and collective memory formation are closely intertwined. Contemporary collective memory of traumatic events have been built also with the intervention of the press, television, radio, cinema and new media. One example are the images of the airplanes crashing into the World Trade Center on September 11 2001, which at the time were broadcast live by the mass media and are now reported in the associated Wikipedia article, as well as engraved in people's minds.

However, the quantitative analysis of edit activity patterns is only a first step towards the study of collective memory processes in Wikipedia. To better

comprehend how traumatic events are represented in the online encyclopedia, and the psychological processes underlying the initial formation of their memories, it is necessary to look more deeply into the content of users' edits and comments. The next chapters will apply automated natural language processing techniques to study the representation of different kinds of events such as traumatic and non traumatic events, natural and human-made disasters, and the initial formation of the collective memory of the London bombings of 7 July 2005 and of the 2011–2012 Egyptian revolution in the associated talk pages.

Chapter 4

The automated natural language processing of events in Wikipedia

4.1 The automated analysis of language and the Linguistic Inquiry and Word Count

The language we use in our daily life reflects our identities, our inner thoughts, fears, desires, and emotions (Tausczik & Pennebaker, 2010). At the individual level, it is our primary form of communication, through which we translate our internal thoughts and establish social relationships with others. At the cultural level, it expresses the collective needs, beliefs and memories of a society. We translate events into words, and the language we use affects the way those events will be later perceived, understood, and recalled, enriching them with moral values and profound meanings, contributing to define the collective identity of a community. Indeed, talking about an event is a sensemaking activity, which acts also as a form of rehearsal, helping to better organize thoughts and memories about the facts (Pennebaker & Harber, 1993).

When people face an emotional upheaval they naturally tend to talk about it, and the social sharing of their feelings serves different psychological purposes, such as making sense of the trauma, seeking for affective and social support, or strengthening the emotional bonding with other members of the community (Arthur, 2008; Cohn et al., 2004; Frijda, 1997; Micalizzi, 2012; Recumber, 2012; Rimé & Christophe, 1997; Wang, 2008). Pennebaker and Harber (1993) studied the degree of social sharing of two traumatic events: the San Francisco Bay Area earthquake of 1989 and the Persian Gulf War. During a period of three months, they interviewed residents of San Francisco and of Dallas, Texas, asking, among other questions, how frequently in the previous 24 hours they had talked with someone about the earthquake or the war. They found that the frequency of social sharing of these traumatic events was considerably higher during the two weeks immediately following the earthquake or the beginning of the war. According to the authors, and considering also that during the first days the media coverage had been intense, this initial period of social sharing provided the basis for the construction of common experiences and collective memories. Talking about an emotional upheaval helps people to better organize and assimilate facts. When an event is particularly meaningful from an emotional point of view, it determines talking and social sharing in the attempt to understand and make sense of it. In this context, social sharing can also have a therapeutic effect, fostering healing and allowing people to move past the traumatizing experience (Arthur, 2008; Micalizzi, 2012; Recumber, 2012; Wang,

2008). Once an emotional upheaval has been cognitively processed and assimilated, people may also forget the trauma (Pennebaker, Paez, et al., 1997).

Typically, past research on traumatic events has implemented interviews and retrospective self-reports to study the psychological consequences of disasters (Cohn et al., 2004), but it has been shown that self-reports can lead to biases and memory distortions (A. A. Stone et al., 2000). Now, with the development of computers, Internet and Web 2.0 environments, we are provided with the unprecedented opportunity to study people's reactions to upheavals as they naturally unfold. The widespread accessibility to people's thoughts and emotions in digital environments provides researchers with new opportunities toward empirical and quantitative work on a large scale. By analyzing word usage patterns during the initial period of social sharing in the aftermath of a traumatic event, it is possible to link the language used to the emotional, social and cognitive processes taking place in the early stage of collective memory formation (Pennebaker & Banasik, 1997; Pentzold, 2009; Tausczik & Pennebaker, 2010). The Internet provides a large amount of data which researchers can collect unobtrusively and almost in real time, and the massive backup into digital archives allows scholars to conduct longitudinal studies on these data, without giving up the spontaneity of interactions (Cohn et al., 2004; Garde-Hanse et al., 2009; Keegan, 2011). In this context, Wikipedia, where users share their thoughts and perspectives to reach a common view on events, seems particularly appropriated for the study of traumatic events (Ferron & Massa, 2011b; Keegan, 2011; Keegan, Gergle, & Contractor, 2011; Pentzold, 2009).

One of the most widely used tools for computational linguistics and natural language processing is WordNet (Fellbaum, 1998). WordNet is a large online lexical database in which different syntactic categories such as nouns, verbs, adjectives and adverbs are organized into sets of synonyms (*synsets*) representing concepts and linked together by conceptual-semantic and lexical relations. Synonymy is the basic relation in WordNet, a simmetric relation between word forms used to represent word senses. Synsets are interlinked to form a navigable network of meaningfully related words and concepts, which can be exploited for performing sophisticated reasoning on language use. WordNet has been used for a variety of purposes, such as word sense disambiguation (Agirre & Martinez, 2000; Agirre & Rigau, 1996), information retrieval (Langer & Hickey, 1998; Magnini & Strapparava, 2001; Mandala, Tokunaga, & Tanaka, 1998), machine translation (Dorr & Katsova, 1998; Knight, 1993), and named entity recognition (Magnini, Negri, Prevete, & Tanev, 2002).

Starting from WordNet, several extensions have been developed for extracting affective and emotional information from text, in the frame of sentiment analysis and opinion mining research. For example, Strapparava and Valitutti (2004) developed WordNet-Affect from WordNet Domains (Magnini & Cavaglià, 2000), a multilingual lexical resource augmenting WordNet with domain labels hierarchically organized (e.g., sport, politics, medicine, architecture, transport). Specifically, the purpose of WordNet-Affect was to include an additional hierarchical structure of affective labels, annotating synsets representing affective concepts. To develop WordNet-Affect, the authors applied affective labels to a number of synsets in order to better specify their affective meaning. For instance, concepts representing emotional states were labeled as "emotion". Other labels identify concepts representing moods (e.g., amiable), traits (e.g., aggressiveness), cognitive (e.g., confusion) or physical states (e.g., illness), edonic signals (e.g., suffering), emotion-eliciting situations (e.g.,

awkwardness), emotional responses (e.g., tremble), behaviors (e.g., offense), attitudes (e.g., intolerance), and sensations (e.g., coldness). WordNet-Affect was later extended including additional categories to distinguish synsets according to their emotional valence, such as positive, negative, ambiguous, or neutral (Strapparava, Valitutti, & Stock, 2006).

Another lexical resource for sentiment analysis and opinion mining developed from WordNet is SentiWordNet (Baccianella & Sebastiani, 2010; Esuli & Sebastiani, 2006), created from the automatic annotation of WordNet synsets according to their degrees of positivity, negativity or neutrality. SentiWordNet was developed to identify, through numerical scores, if and to what extent a term is a marker of opinionated, subjective content, or if it is objective. The program allows to extract opinions from texts, identifying whether a subjective word representing an opinionated content is positively or negatively connoted.

Similarly, Q-WordNet, developed by Agerri and Garcia-Serrano (2010), is a lexical resource for the polarity classification of terms, in which WordNet senses are automatically annotated by positive and negative polarity. Determining whether a text is positively or negatively connoted is central to opinion mining research, and applications for polarity classification can be employed for the identification of opinions about commercial products and brand monitoring, but also for the study of attitudes on online forums, blogs and other digital environments.

A different perspective to the study of language is the word count approach, according to which word use can be quantified and statistically analyzed. This methodology, in the frame of quantitative text analysis, has been used for the study of both content and language style (Pennebaker, Mehl, & Niederhoffer, 2003; Tausczik & Pennebaker, 2010). Current approaches involve simple word count, for example of grammatical units such as prepositions and personal pronouns, or of psychological related linguistic dimensions, such as words indicating emotions, regardless of their semantic context.

Gottschalk and colleagues (Gottschalk & Gleser, 1969; Gottschalk, Gleser, Daniels, & Block, 1958; Gottschalk, Winget, & Gleser, 1969) developed a content analysis method to measure mental and emotional states derivable from verbal behavior. Transcriptions of patients' speech samples were evaluated by judges to identify language aspects related to social and psychological topics, such as anxiety or hostility toward others. The Gottschalk-Gleser method was applied to the psychiatric diagnoses of cognitive impairments and mental disorders such as anxiety, dependency or social alienation (Bantum & Owen, 2009; Tausczik & Pennebaker, 2010). The scales were later employed in the development of the computer program Psychiatric Content Analysis and Diagnosis (PCAD; Gottschalk & Bechtel, 1993) which, despite having established an adequate construct validity and reliability (Gottschalk, 1995), does not provide clear scoring rules for detecting emotions in texts (Bantum & Owen, 2009).

Philip Stone and colleagues developed General Inquirer, the first general purpose computer-assisted approach for content analysis of textual data in psychology (Rosenberg & Tucker, 1978; P. J. Stone, Dunphy, Smith, & Ogilvie, 1966), which maps text files with counts on dictionary-supplied content analysis categories consisting of lists of words and word senses. General Inquirer employs disambiguation routines and has proven to be effective in the evaluation of speech to identify mental disorders and personality traits. However, one limitation of its earlier versions is that they were based on the manipulation and weighting of language

dimension chosen by the category developers and not visible to users (Tausczik & Pennebaker, 2010).

Hart (1984, 2001) created DICTION, a program designed to count words in political messages. The program detects the frequency of use of five categories, each one divided into more specific classes. The main categories, specifically conceived in the framework of political discourse analysis, are: certainty, indicating resoluteness and inflexibility; optimism, related to positive thinking; activity, associated to references to movement, change and new ideas; realism, highlighting tangible matters affecting people's everyday lives; and, finally, commonality, indicating group and collective values.

In order to provide a tool to support the analysis of psychotherapy sessions, Mergenthaler and colleagues developed TAS/C (Mergenthaler, 1996), which focuses on two language dimensions: emotional tone, related to the density of emotional words in a text, and abstraction, defined as the frequency of abstract nouns, identified through the use of suffixes such as "-ity", "-ness", "-ment". Subsequently, TAS/C was extended to measure referential activity, indicating the ability to verbalize different kinds of nonverbal experiences, including imagery and emotional experiences (Bucci, 1995). TAS/C has been successfully applied to the analysis of therapy protocols, proving to be a useful tool for detecting clinically important moments in transcripts from psychoanalytic sessions (Mergenthaler, 1996).

Asking participants to speak into a microphone for 10 minutes, Weintraub (1981, 1989) developed a methodology for sampling natural language, with a specific orientation toward psychoanalysis and psychotherapy. Unlike other word count approaches using computer programs, Weintraub linguistic analysis employed naive judges with limited knowledge of linguistics, who rated the texts on several dimensions drawn from clinical experience with patients with psychopathological problems. The linguistic dimensions considered by Weintraub included pronoun categories, (e.g., I, we), negatives (e.g., not, never), qualifiers (e.g., kind of), feelings (e.g., I love), and adverbial intensifiers (e.g., really).

One of the tools successfully employed for studying the emotional, cognitive and structural components of written text related to traumatic upheavals is the text analysis application Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2001). LIWC was designed to calculate the degree to which people use different categories of words in a text, assigning words to various linguistic and psychological dimensions of language. LIWC has two main features: a processing component, which analyzes each text file on a word-by-word basis, and a dictionary, which is a collection of words defining particular categories. Specifically, LIWC reads a text file analyzing one word at a time, looking for a dictionary match with each target word in the text. If the current word matches a dictionary word, the corresponding word category scale is incremented. For example, the word "cried" is assigned to the categories "verbs", "past tense", "sadness", "negative emotions", and "affective processes". Hence, if it is found in the analyzed text, each of these categories scale scores will be incremented. After processing all the words in a text file, LIWC calculates the percentage of words for each category. Categories in the dictionary are generally organized hierarchically (see *Table 4.1*): for example, all words included in the category "positive emotions" are also included in the overarching category "affective processes", which is the sum of "positive emotions" and "negative emotions" (Pennebaker et al., 2001).

The simplicity of LIWC's processing strategy was found to perform well in general, but the fact that the program allows for minimal contextual analysis remains a limitation, making it difficult to distinguish between different meanings of the same word (Chung & Pennebaker, 2007). Indeed, while language dimensions are often straightforward, other psychological categories are more subjective and context dependent. Bantum and Owen (2009) analyzed emotional expressions in a web-based intervention for women with breast cancer using rater coding, LIWC and PCAD (Gottschalk & Gleser, 1969; Gottschalk et al., 1958; Gottschalk et al., 1969), developed to assist clinicians with therapy sessions and psychiatric diagnoses. Although both programs were found to over-identify emotional expressions, authors suggested that LIWC performed better that PCAD for the identification of emotions in a text. Specifically, LIWC showed a stronger performance in general emotion categories (i.e., affective processes) than in specific types of emotions (e.g., anxiety), and was found to be 12 times more likely to over-identify than to under-identify emotional words. For example, some words that were frequently coded as emotions by LIWC but not by raters were "good", "hope", "beautiful", and "best". Nevertheless, LIWC allows for limited contextual analysis for particular words, such as "like" and "kind". For example, the word "like" is assigned to the category "affective processes" and "positive emotions" if it is preceded by a pronoun or by a word indicating "discrepancy" (e.g., "would"), otherwise it is categorized as a "filler" (e.g., "youknow"). Although LIWC's accuracy could be certainly improved employing more sophisticated computational strategies for word disambiguation (Agirre & Edmonds, 2006; Weik, 1996), it has proven to perform well for emotion detection, and its content and construct validity was successfully assessed by Pennebaker and colleagues (Pennebaker & Francis, 1996; Pennebaker et al., 2001). Furthermore, Bantum and Owen (2009) found LIWC to have a good convergent and discriminant validity for the analysis of emotional content in a text.

One of the major criticisms moved to LIWC and similar software is that, since communication is complex and multilayered, language must be considered within a specific context (Mehl, 2006; Pennebaker et al., 2003). The fact that most word count approaches are blind to context is certainly a major shortcoming. For example, they are unable to recognize irony, sarcasm or metaphoric language, and they cannot identify different meanings of the same words in different contexts. However, in the past years word count approaches have proved their validity in different domains of psychology (see for instance Gottschalk, 1995; Pennebaker et al., 2003; P. J. Stone et al., 1966; Weintraub, 1981). Furthermore, from a measurement perspective, there is evidence suggesting good psychometric properties of word count based measures (Mehl, 2006).

In this work, the text analysis application Linguistic Inquiry and Word Count (Pennebaker et al., 2001) will be applied to the study of the collective representation of traumatic events in Wikipedia pages. Three main reasons motivated this choice. First, previous research employed it for investigating the effects of traumatic events (Back et al., 2010; Cohn et al., 2004; Keegan, 2011), and this allows the comparison of results across different studies. Second, its fine-grained analysis of language related to different psychological processes allowed to analyze specific consequences of collective upheavals on affective, social and cognitive processes, which were previously investigated in theoretical and empirical research (for example, Baum, 1987; Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Mehl & Pennebaker, 2003; Pennebaker et al., 2003; Pyszczynski et al., 2003; Rimé et al., 1998;

Stroebe, Hansson, Stroebe, & Schut, 2001). Third, despite their substantial limitations, LIWC and other word count approaches have the pragmatic advantage of providing powerful and easy-to-use tools for psychological research, without requiring preformatting, text parsing and data annotation such as more sophisticated tools (Mehl, 2006). In other words, in its simplicity, despite being quite superficial and error prone, LIWC can be an effective tool for the study of the psychological processes underlying the collective memory building of traumatic events.

This chapter aims at assessing the effectiveness of LIWC for the empirical and quantitative analysis of the language used in Wikipedia pages about traumatic events. The following sections will apply LIWC to Wikipedia articles, to assess whether it can be employed to detect specific psychological processes underlying different patterns of language, providing the basis for the implementation of automated content analysis techniques for the study of collective memory processes on Wikipedia.

4.2 Linguistic markers of traumatic and non traumatic events

A first step toward the study of the representation of traumatic events in Wikipedia was to test if LIWC can be applied to English Wikipedia articles for detecting meaningful differences in the use of words related to the main psychological categories between different kinds of pages. For this reason, we first compared the content of a sample of articles related to traumatic events (bombings, assassinations, earthquakes, etc.) to a sample of articles about non traumatic events (such as music events, royal weddings, sport competitions).

Past research on the psychological consequences of traumatic events found a temporary growth in the negative mood after an emotional upheaval (Koss & Kilpatrick, 2001; Stroebe et al., 2001), an increase in cognitive processing as sensemaking activities take place and people try to comprehend and eventually find a meaning in what happened (Davis & Nolen-Hoeksema, 2001; Pennebaker et al., 2003), and an increase in social sharing and social interactions (Mehl & Pennebaker, 2003; Pyszczynski et al., 2003; Rimé et al., 1998). This section focuses particularly on affective, cognitive and social processes, represented in LIWC by the corresponding categories and subcategories reported in *Table 4.1*. Specifically, in the comparison between articles about traumatic and non traumatic events, we expected to find higher amounts of words related to negative emotions, cognitive and social processes in the former, and a higher presence of language related to positive emotions in the latter.

4.2.1 Method

Using Wikipedia internal categories such as "Events by topic", Britannica Online Encyclopedia (2011), History Central (2011), and Information Britain (2011) we identified 66 articles related to traumatic events, such as "September 11 attacks" or "2004 Indian Ocean earthquake and tsunami". Through Wikipedia internal categories such as "Sports events", "List of most watched sporting events in 2004", "Largest concerts ever", "Music events" we identified 40 articles about non traumatic events, like "Coronation of Queen Elizabeth II" or "2010 FIFA World Cup". The complete

list of articles, along with their LIWC scores is available at http://sonetlab.fbk.eu/data.

LIWC was applied to the text of each article to obtain a score for the psychological categories of theoretical interest that operationalized our hypotheses (*Table 4.1*), and all measures were computed as percentages of words in relation to the total number of words contained in each text.

Psychological processes	Examples
Social processes	Mate, guy, boy
Family	Daughter, brother, dad
Friends	Buddy, friend, mate
Humans	Adult, children, girl
Affective processes	Happy, hate, kiss
Positive emotions	Love, party, pleasant
Negative emotions	Hurt, abuse, scary
Anxiety	Worried, afraid, apprehensive
Anger	Kill, aggression, destroy
Sadness	Sad, cry, depression
Cognitive processes	Cause, acknowledge, admit
Insight	Think, assume, interpret
Causation	Because, depend, elicit
Discrepancy	Should, could, if
Tentative	Maybe, apparently, suppose
Certainty	Always, absolutely, clear
Inhibition	Block, abstain, avoid
Inclusive	And, add, along
Exclusive	But, either, without

Table 4.1 LIWC psychological categories of primary interest to the study and their hierarchical categorization.

In order to convert percentages into more centered scores, we applied the arcsine transformation, traditionally employed in general linear modeling to analyze a dependent variable when the raw values are proportions or percentages. This transformation consists in taking the arcsine of the square root of a number and transforming the original data in radians, allowing skewed distributions to better approximate normality (Osborne, 2010; Sokal, 1981).

After assessing for the assumption of normality with the Kolmogorov-Smirnov test, a series of independent samples t-tests were applied to compare the scores of the different psychological variables between articles related to traumatic and non traumatic events. Where the assumption of normality was violated, the Mann-Whitney U non parametric test was applied (*Figure 4.1*).

4.2.2 Results

In general, the results showed a significantly higher presence of language related to affective processes ($t_{(104)}=3.90$, $p<.001$), negative emotions ($U=64$, $p<.001$), cognitive

(U=983.5, p=.028) and social processes for articles about traumatic events. *Figure 4.1* shows the results for the overarching categories and the subcategories of interest. Specifically, the relative number of words expressing anxiety (e.g., "worried"), anger (e.g., "hate") and sadness (e.g., "cry") was significantly higher in articles about traumatic events (respectively, $t_{(104)}$=6.47, p<.001; $t_{(102.96)}$=9.44, p<.001; $t_{(104)}$=3.94, p<.001), while language associated with positive emotions was significantly higher in articles about non traumatic events ($t_{(68.56)}$=-6.76, p<.001).

The amount of words related to cognitive activity, such as those expressing insight, causation and tentative language (e.g., "think", "because", "maybe") was found to be significantly higher in articles about traumatic events (respectively, $t_{(104)}$=4.73, p<.001; $t_{(104)}$=2.79, p=.006; $t_{(104)}$=3.58, p=.001), confirming the presence of sensemaking activities when creating the collective memory of emotional upheavals (Pennebaker et al., 2003).

With regard to social processes, the t-tests showed a significantly higher amount of words expressing references to other people in general, such as humans (e.g., "adult") in articles about traumatic events ($t_{(104)}$=4.38, p<.001). In the same direction, the Mann-Whitney U test showed a significantly higher presence of references to family members (e.g., "daughter), with U=971.5, p=.021. On the contrary, probably reflecting a more informal and relaxed content, references to friends (e.g., "buddy") were found to be significantly higher in articles about non traumatic events (U=935.5, p=.011). All other differences for the remaining psychological categories reported in *Table 4.1* did not attain statistical significance and therefore are not represented in *Figure 4.1*.
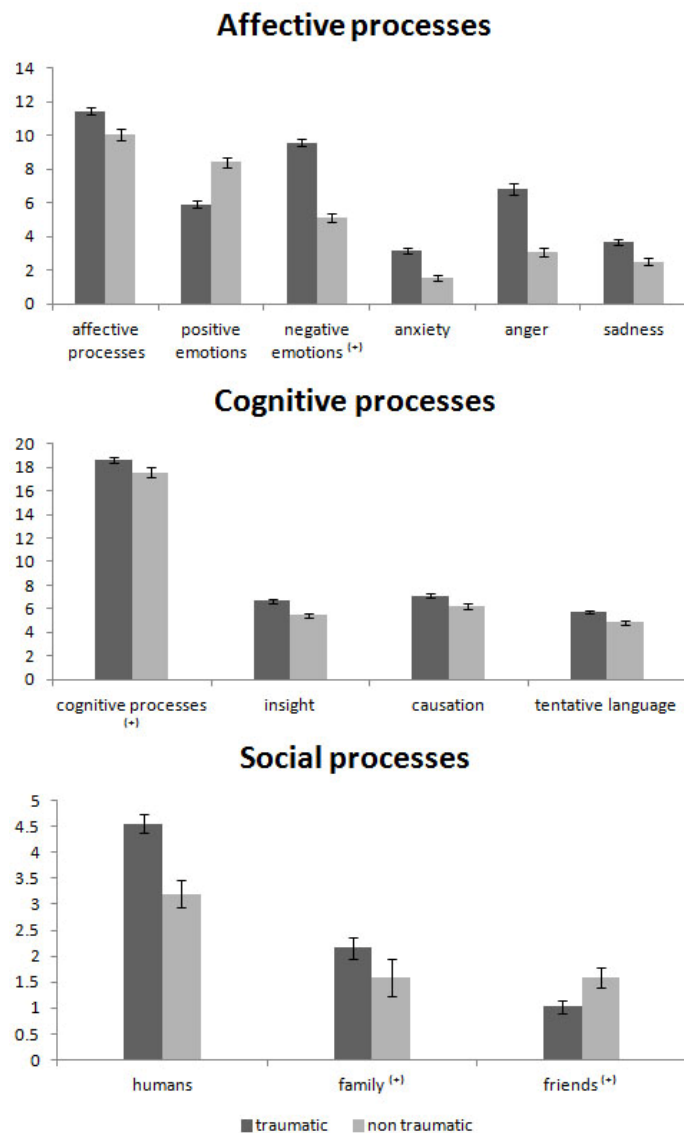
Affective processes

Cognitive processes

Social processes

Figure 4.1 Arcsine percentages of words related to different psychological categories for articles about traumatic and non traumatic events. The symbol [+] indicates variables for which the assumption of normality had been violated, and whose scores were compared through Mann-Whitney U non parametrical test. All differences are statistically significant.

Despite not being particularly insightful nor surprising results, the generalized higher percentages of language related to affective, cognitive and social activity in articles about traumatic events confirmed the presence of specific psychological processes traditionally related to emotional upheavals (Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Mehl & Pennebaker, 2003; Pennebaker et al., 2003; Pyszczynski et al., 2003; Rimé et al., 1998; Stroebe et al., 2001). LIWC ability to detect significant differences in the expected language patterns of very dissimilar types of articles confirms it can be reasonably employed for further analyses of more specific processes characterizing the collective representation of traumatic events in Wikipedia.

4.3 Temporal focus of recent and old traumatic events

Given their encyclopedic nature, Wikipedia articles about historical events should be generally written in the past tense, and breaking news should not be treated differently from other events, nor be written in news style ("Wikipedia:Writing better articles", 2004; "Wikipedia:What Wikipedia is not", 2001). In Wikipedia, breaking news receive a considerable attention by users. Indeed, pages about recent traumatic events tend to get created just few hours or days after their happening (Ferron & Massa, 2011b; Keegan, 2011).

(Keegan, 2011) selected 127 Wikipedia articles listed under the "List of accidents and incidents involving commercial aircraft", classifying them as breaking articles, started within 48 hours of the crash, and non-breaking articles, written more than two days after the accident. After processing every revision with LIWC in order to study the psychological states of users derivable from their edits to the articles, he found a non significant decrease in the narrative immediacy, showing that, compared to breaking articles, non-breaking articles have an accelerating tendency to contain more past-tense than present-tense verbs over time.

At the early stage of articles related to breaking news, a higher use of present tense rather than past tense verbs might reflect that sensemaking processes are ongoing and editors are still trying to understand the facts, organize the sourcing and interpret the events (Keegan, 2011). Analyzing the tense of common verbs employed in Wikipedia articles can tell us more about the temporal focus of their editors (Tausczik & Pennebaker, 2010). Furthermore, a distinct presence of past, present and future tenses at the initial stage of articles about old and recent traumatic events could empirically support the distinction theorized by Assmann (1995) between communicative and cultural memory, which Pentzold (2009) suggested to apply to Wikipedia in its interpretation as a global memory place. According to Assmann (1995), collective memories form through the gradual passage from communicative memory, highly interactive, unstable, based on everyday communication and with a limited temporal horizon, to a cultural memory, more formal, fixed, objective and distant from everyday interactions.

For these reasons, in the analysis of Wikipedia articles about emotional upheavals at their early stage, we expected a higher presence of past tense verbs in those about old traumatic events, happened before Wikipedia was created in 2001, and therefore crystallized in the collective memory of individuals. On the contrary, we expected a higher presence of present and future tenses in articles about recent traumatic events, indicating that the construction of collective memories was still at a communicative phase.

4.3.1 Method

This section focuses specifically on traumatic events by analyzing the differences between those happened before Wikipedia was created in 2001 (old traumatic events) and those occurred after this threshold (recent traumatic events). In order to consider the formation of collective memory at its early stage, articles were analyzed at their 500th edit. This threshold allowed to restrict the dataset to a group of articles with a sufficiently large amount of text, without removing too many pages. At the same time, considering early versions, the analysis could focus on the process of

collective memory creation in its beginning (Assmann, 1995). Requiring to have at least 500 edits reduced the dataset of traumatic events to 55 Wikipedia articles, 26 of which were old traumatic events and 29 were recent accidents and disasters.

As in the previous analysis, LIWC was used on each text to obtain scores for linguistic categories about the tenses of common verbs. The measures were computed as percentages of words in relation to the total number of words contained in each text. The arcsine transformation was applied, and after testing for normality with the Kolmogorov-Smirnov test, the scores of past, present and future tenses in articles about recent and old traumatic events were compared through independent samples t-tests (*Figure 4.2*).

4.3.2 Results

The graph in *Figure 4.2* shows a significantly higher presence of past tense verbs in articles about traumatic events happened before 2001 ($t_{(53)}$=2.59, p=0.12), and a significantly higher amount of present and future tense verbs in articles related to more recent events (respectively, $t_{(43.67)}$=-3.64, p=.001; and $t_{(53)}$=-2.56, p=.013).
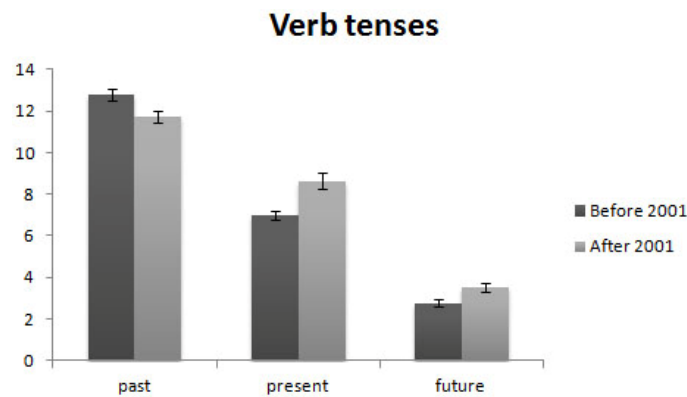

Figure 4.2 Arcsine percentages of past, present and future tense verbs in articles about recent and old traumatic events. All differences are statistically significant.

This distinct presence of verb tenses in articles about old and recent traumatizing episodes confirmed that Wikipedia editors of articles about breaking news, at least at a early stage, were still focused in the present, despite Wikipedia guidelines ("Wikipedia:Writing better articles", 2004; "Wikipedia:What Wikipedia is not", 2001). On the contrary, articles about traumatic events happened before the creation of Wikipedia contained more past tense verbs, suggesting more stable and objective contents.

4.4 Discussion

In summary, this chapter showed that LIWC can be effectively applied to detect different types of psychological processes in articles about traumatic and non traumatic events, and a different use of common verb tenses in articles about recent and old traumatic events. In particular, its ability to identify an expected distinctive presence of language related to specific psychological processes in articles about very

different topics such as traumatic and non traumatic events confirmed LIWC as an efficient tool for the analysis of collective memory building processes emerging from Wikipedia.

The higher presence of words related to negative emotions, cognitive activity and social processes in articles about traumatic events was not surprising but highlighted expected characteristics of their collective representation. Negatively connoted affective language reflected the nature of traumatizing events, which are strongly emotionally charged, while the presence of language suggesting an increased cognitive activity in articles about emotional upheavals reflected the need to understand the reasons and find reasonable explanations in what happened.

Specifically, the use of insight (e.g., think, assume, interpret), and causal words (e.g., because, depend, elicit) in the description of traumatic events suggested an active sensemaking process of appraisal and reasoning. In past research, the presence of insight and causal words in expressive writings was found to be related to better health improvements after a traumatic experience (Pennebaker, Mayne, & Francis, 1997). Similarly, the increased use of tentative language (e.g., maybe, apparently, suppose) in describing traumatic events reflected the lack of certainty and the sense of insecurity typical of most traumatic events, and at the same time highlighted users' efforts in finding plausible explanations.

With regard to social processes, the higher presence of words related to human beings (e.g., adult, children, girl) and to family (e.g., daughter, brother, dad) confirmed the increased social orientation characterizing the aftermath of traumatic events (Cohn et al., 2004; Pennebaker & Harber, 1993; Rimé & Christophe, 1997), while the increased presence of words related to positive emotions (e.g., love, party, pleasant) and to the friendship domain (e.g., buddy, friend, mate) in articles about non traumatic events reflected a more informal, untroubled and placid context.

The results on the different temporal focus of articles about old and recent traumatic events were consistent with previous research by Keegan (2011), who found an accelerating tendency of non-breaking articles to contain more past-tense than present-tense verbs over time. Most importantly, despite Wikipedia policies suggesting to use an historical perspective and to avoid news style ("Wikipedia:Writing better articles", 2004; "Wikipedia:What Wikipedia is not", 2001), the increased use of present and future tenses in articles about breaking news confirmed Assmann's (1995) theoretical distinction between communicative and cultural memory. This suggests that the initial collective representation of recent traumatic events in Wikipedia was still at the communicative phase of memory building, that is unstable, highly interactive and located in the present. On the contrary, the collective memories of older traumatic events had the time to stabilize and consolidate, and their representation reflected their objectivity and distance from the present in an increased use of past tenses.

In summary, despite the limitations of word count approaches, these results confirmed LIWC as an efficient tool for the automated analysis of Wikipedia articles, allowing to go one step further in the study of the collective representation of traumatic events in Wikipedia, as explained in the next chapters.

Chapter 5

Natural and human-made traumatic events in Wikipedia

5.1 The effects of natural and human-made disasters

This chapter focuses on English Wikipedia articles about natural and human-made disasters, investigating the different representations of these traumatic events and highlighting the diverse psychological processes underlying their shared accounts, exploiting automated content analysis techniques.

Previous studies investigated the effects of single traumatizing events (Adler, 1943; Cohn et al., 2004; Gleser et al., 1981; Gleser et al., 1978; Green, 1980; Henderson & Bostock, 1977; Leopold & Dillon, 1963; Ploeger, 1977; Titchener & Kapp, 1976) but few of them compared natural and human-made disasters, partly because each one is unique in its consequences on the psychological and physical health of the involved populations (Green, 1996). Outlining common effects of different kinds of traumatic events can be difficult, because of the peculiarity of their particular circumstances, the number of dead and injured people, the damage to properties, the exposure to toxic chemicals or to the death of others, the degree of life threatening experiences, etc. Moreover, each individual can react in a different way, and the recovery environments can be very different (Green, 1996).

However, past research suggested theoretical reasons for hypothesizing substantial differences in the type, severity and longevity of consequences of natural and human-made disasters (Baum, 1987; Gleser et al., 1981). In this chapter LIWC was applied to the English Wikipedia to quantitatively investigate the differences in the language used in articles to describe natural and man-made disasters, reporting how the language used can reflect different underlying psychological processes (Tausczik & Pennebaker, 2010), and highlighting the differential nature of traumas.

Understanding how people react to collective emotional upheavals is crucial to a better comprehension of the consequences on the physical and psychological health of many involved communities (Baum, 1987). In fact, for a long time past research has been focusing on the outcomes of specific individual disasters. For instance, Adler's study of the effects of the 1942 Cocoanut Grove fire (Adler, 1943) showed a prolonged persistence of nervousness and anxiety up to 11 months after the fire, while Green (1980) found evidence of emotional problems up to 15 months after another nightclub fire at Beverly Hills. Other studies investigated longer-term effects of traumatizing events and found an increase of psychiatric, psychological, and work-related problems up to several years after a marine explosion, and other psychological problems up to 10 years after a mine cave-in (Henderson & Bostock, 1977; Leopold & Dillon, 1963; Ploeger, 1977). In their investigation on the

consequences of the dam collapse and flood at Buffalo Creek in West Virginia, Titchener and Kapp (1976) reported high rates of emotional problems like anxiety, depression and personality changes, while other studies showed also evidence of hostility, in addition to sleep disturbances and psychiatric problems up to more than 2 years after the accident (Gleser et al., 1981; Gleser et al., 1978). Pennebaker and Harber (Pennebaker & Harber, 1993) studied the aftermath of the Loma Prieta Erthquake, which shocked the San Francisco Bay area in 1989, and found an increase of self-reported illness episodes, quake-related dreams, arguments with family members and co-workers, and also aggravated assaults, in the first weeks following the earthquake.

As already mentioned in Chapter 4, earlier research showed that immediately after a traumatic event nearby residents tend to talk more about the accident, and this increased frequency of social sharing can provide the basis for the future construction of collective memories (Pennebaker, 1997a; Pennebaker & Harber, 1993). With time these collective memories, conceived starting from the immediate responses to disasters, will influence the cultural attitudes toward the current society, directing present and future behavior to ensure the social cohesion and the continuity of the community, for instance engendering protective responses after a threatening traumatic event (Pennebaker, Paez, et al., 1997; Pyszczynski et al., 2003; Reser, 2007). But what kind of memories are these? Is it possible to detect meaningful characteristics of traumatizing events associated to different psychological responses in the aftermath of these upheavals and different types of collective narratives in Wikipedia?

Clearly, traumatic events can be studied from several perspectives and levels of analysis. Disasters can highly differ in nature: they can be meteorological or climatic events, technological mishaps, biological hazards, political or economic crises, all posing potentially serious consequences on both the social and psychological levels, like collective migration, social disruption, loss of property or life, trauma, distress and shock. Of course, this variability makes disasters hard to define. As a matter of fact, it is the precise combination of different characteristics – like the nature of the event, its impact and the victims' responses – that makes each disaster unique (Green, 1996).

In this diversified framework, identifying the conditions leading to particular effects and social responses becomes crucial to understand the various psychological implications of traumatic events, and eventually to provide a basis for the prediction of short and long term consequences (Baum, 1987). The analysis of collective emotional upheavals in terms of their characteristics, such as natural or man-made traumatic events, can provide researchers with additional tools to predict the extent of trauma, the consequences on individuals' psychological health and on communities' social cohesion.

Of course, there are numerous variables mediating disasters' responses, and this makes it is difficult to compare natural and human-made disasters, partly because each emotional upheaval is unique, and so are the subsequent consequences (Green, 1996). It is challenging to assess common outcomes of disasters of a certain kind because there is a considerable variability in the nature of the physical agent, its impact and the emotional responses it causes. Each individual can react in a different way, the exposure to life-threatening situations may vary among the population, and the recovery environments can be very dissimilar. Notwithstanding, there are

theoretical reasons for hypothesizing different outcomes of natural and man-made catastrophes (Green, 1996).

Andrew Baum (Baum, 1987), in reviewing previous research on natural and man-made disasters, proposed that accidents caused by man may have more important effects in terms of longevity and severity. Starting from the analysis of the accident at Three Mile Island, which despite causing little physical damage to the population had a long aftermath of more than three years, Baum and colleagues (Baum, Fleming, & Reddy, 1986) proposed a characterization of natural and technological disasters in terms of their characteristics:

Suddenness. Although there is still variability in the suddenness of natural and man-made disasters, in general both can have a rapid onset. While a tornado or a storm can take days to form, an earthquake can give little warning to the population. Industrial explosions or air crashes are usually sudden and unexpected. There are also human-made disasters with a slow onset, like heat waves, the poisoning of the oceans or economic crisis, but usually they are not considered as traumatic events (Reser, 2007).

Power. Both natural and technological disasters can be highly powerful and threatening, causing death and destruction.

Visible damage. While generally natural disasters cause damage to the environment and destruction of properties, this is not always true for technological disasters, as the Three Mile Island or the Chernobyl accidents, where invisible damage is related to illness in the long-term.

Predictability. Thanks to the improvement of forecasting abilities, some natural disasters can now be prevented to a certain extent (e.g., storms or tornadoes). On the contrary, technological disasters are hardly predictable, because technology is not expected to fail. They are also usually sudden, leaving little or no time for evacuation.

Low point. Natural disasters usually have a clear and identifiable low point, defined as the moment when the worst has already happened and after which the focus can be moved to recovery efforts; for some technological catastrophes, like those involving radiation or toxic leaks, it is more difficult to identify a specific low point.

Perception of control. Natural disasters are usually perceived as uncontrollable. There may be political controversies related to the management of the disaster, but they mainly depend on our inability to control the natural elements. On the contrary, technology itself is the manifestation of human control over the environment, so when it fails, the disaster is perceived as a loss of control. Losing control, meaning not having control when there are expectations for it, seems to cause different psycho-physiological consequences than not having control on an unexpected accident (Baum & Gatchel, 1981; Wortman & Brehm, 1975). Baum and colleagues (Baum et al., 1986) showed that in the former case the loss of control can be related to stress arousal, while in the latter, when control is not expected, people's reactions tend to helplessness and passive behavior.

Extent of effects. In the case of natural disasters, effects are usually bounded to the people directly involved and to a limited area around the accident; for technological disasters, the effects may be broader and involve also people not directly affected by the accident, due to a general loss of confidence and credibility.

Persistence of effects. According to Baum and colleagues (Baum, Fleming, & Davidson, 1983), it seems that for natural disasters the effects may be limited to the short-term, even if their extent depends upon the individual experience of the

trauma, such as the loss of property or of loved ones; with regard to technological disasters, the effects seem to be more prolonged in time, particularly for toxic accidents.

This chapter focuses on the differences between the collective representations of natural and human-made traumatic events analyzing the content of the related English Wikipedia articles, and particularly the language associated with affective, cognitive and social processes. As already shown in Chapter 4, articles related to traumatic events generally contain higher amounts of words reflecting negative emotions, cognitive and social processes. In the comparison between articles about natural and man-made disasters, drawing from Baum's theoretical distinctions and focusing particularly on the loss of control (Baum et al., 1983), we expected a distinctive presence of words related to anger, anxiety and sadness. Since natural disasters are generally perceived as uncontrollable and therefore may engender more passive responses (Baum et al., 1986) we expected a higher presence of language related to sadness in the associated articles. On the contrary, in case of human-made disasters, the sudden and unexpected loss of control may provoke more focused anger and anxiety. The perception that these kind of accidents can, in theory, be prevented, may amplify the extent of trauma triggering violent outcomes. Especially in case of terrorist accidents, subsequent reactions may find expression in anger and blame because of the purposeful injury of other people.

With regard to cognitive processes, although the search for causes and accountability is a typical consequence also of natural disasters (Reser, 2007), we expected a higher presence of language related to cognitive activity in articles about human-made disasters, because in this case there is an unpredictable loss of control which may require more exhaustive and comprehensive explanations (Baum, 1987).

We also expected a higher amount of words related to social processes in articles about man-made traumatic events, partly because this kind of collective upheavals may affect more deeply people's social lives, triggering an increased orientation toward others (Mehl & Pennebaker, 2003; Pyszczynski et al., 2003; Rimé et al., 1998).

5.2 Method

Exploiting Wikipedia internal categories such as "Events by topic", and other online encyclopedias and digital resources such as Britannica Online Encyclopedia (2011), History Central (2011), and Information Britain (2011), 55 Wikipedia articles about traumatic events were identified. The categorizations made by users through Wikipedia internal lists and categories, like "List of natural disasters by death toll", "2004 natural disasters in the United States", "Man-made disasters", or "Terrorist incidents" were exploited to classify traumatic events as human-made or natural. Articles belonging to "Man-made disasters", "Mass murder", "Assassinations", and other similar categories were classified as human-made disasters, while articles assigned by Wikipedia users to categories like "1993 natural disasters" or "2008 Atlantic hurricane season" were classified as natural disasters (the complete list of articles with their related categories was released at http://sonetlab.fbk.eu/data). As in the study on the temporal focus of old and recent traumatic events described in Chapter 4, the specific interest was on the early versions of articles, to consider the collective representation of events at the communicative level of memory formation

(Assmann, 1995). For this reason articles were considered at their 500th revision, a threshold allowing to analyze pages containing a substantial amount of text. Out of 55 articles, 19 were about natural disasters and 36 were about human-made traumatic events.

LIWC was applied to each article to get a score for the different psychological categories already reported in *Table 4.1*, and all measures were computed as percentages of words in relation to the total number of words contained in each text. The arcsine transformation, explained in Chapter 4, was applied and the assumption of normality was tested with Kolmogorov-Smirnov tests: where it was not violated a series of independent samples t-tests was applied to compare the presence of words related to different categories in the articles about natural or man-made disasters, otherwise Mann-Whitney U non parametric tests were executed (*Figure 5.1*).

5.3 Results

The results reported in *Figure 5.1* confirmed our hypotheses, showing in general significantly higher amounts of emotional language related to anger and anxiety, cognitive and social processes for articles about man-made traumatic events, and more sadness for natural disasters. In particular, t-tests results showed that words expressing anxiety and anger (e.g., "worried", "hate") were significantly higher in articles related to human-made disasters (respectively, $t_{(53)}$=2.79, p=.007; $t_{(53)}$=2.63, p=.011), while on the other hand the amount of words associated to sadness (e.g., "cry") was significantly higher in articles about natural disasters ($t_{(53)}$=-3.812, p<.001). The language expressing cognitive activity, insight (e.g., "because", "think"), inhibition (e.g., "prevent", "forget") and exclusion (e.g., "but", "either") was, as expected, more present in articles about man-made traumatic events (respectively, $t_{(53)}$=3.47, p=.001; $t_{(53)}$=4.51, p<.001; $t_{(53)}$=5.57, p<.001; U=226.5, p=.041). Moreover, the amount of words referring to social processes and to family (e.g., "they", "son") was significantly higher in articles related to human-made disasters (respectively, $t_{(53)}$=4.79, p<.001; U=211, p=.020). No other categories among those reported in *Table 4.1* attained statistical significance.
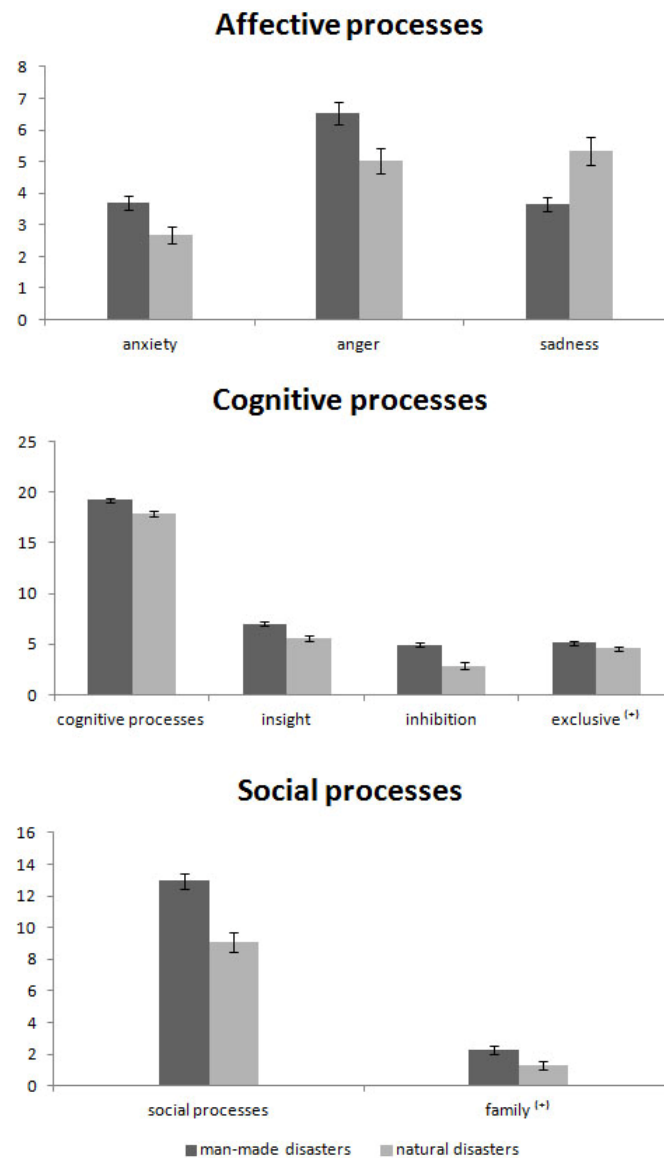
Figure 5.1 Arcsine percentages of words related to different psychological categories for articles about man-made and natural disasters. The symbol [+] indicates variables for which the assumption of normality had been violated, and which scores were compared through Mann-Whitney U non parametrical test. All differences are statistically significant.

These results confirmed that natural and man-made disasters are represented with different language patterns in Wikipedia, suggesting that distinctive psychological and sensemaking processes may underlie users' collaboration to the editing of these articles. In particular, articles about human-made traumatic events were characterized by language referring to anger and anxiety, while the collective representation of natural disasters expressed more sadness. The language patterns associated to different affective processes in these articles support past literature suggesting that traumatic events caused by human agencies may be more stressful, trigger more anxiety and focused anger than natural disasters, which instead may be characterized by more passive responses. Although the urge to understand the causes and find a

convincing explanation is crucial for both these types of disasters, it may be even more critical for man-made traumatic events, where the loss of control is often perceived as a serious damage to the community and the need to blame somebody for failure is more pressing. Cognitive words refer to mental activities involving reflection and critical thinking, and they were found to be significantly higher in articles about human-made traumatic events, with respect to those on natural disasters. In particular, words revealing insightful thinking (such as "realize", "understand", "think", "consider") are indicators of reflection and rationalization processes carried out by people trying to find explanations and work through the events (Pennebaker & Francis, 1996). Language expressing exclusion (e.g., prepositions and conjunctions such as "but", "without", "except", "exclude"), which is helpful in making distinctions and differentiations, has been related to cognitive complexity (Newman, Pennebaker, Berry, & Richards, 2003; Pennebaker et al., 2003). Studying deceptive language, Newman and colleagues (2003) found that liars used fewer exclusive words with respect to truth-tellers. The authors concluded that, since telling false stories and providing information about something which did not happened is cognitively complicated, liars may prefer less complex cognitive strategies. Thus, the presence of significantly higher amounts of exclusive words in articles about human-made disasters suggests an increased cognitive complexity in the construction of coherent narratives of such events (Pennebaker et al., 2003). Moreover, words expressing inhibitory processes (e.g., "avoid", "deny", "forget", "neglect", "refuse"), related to psychological inhibition, were found to be higher in articles about man-made traumatizing events. In past research psychological inhibition, or the suppression of thoughts and feelings, was found to be associated with negative health consequences (McGuire & Kiecolt-Glaser, 2000; Pennebaker, 1997b; Petrie, Booth, & Pennebaker, 1998). Previous studies showed that individuals who were asked to write about personal stressful experiences reported fewer medical visits in the following months compared to a control group who wrote about more trivial topics (Pennebaker, 1993; Pennebaker, Barger, & Tiebout, 1989; Pennebaker & Beall, 1986). Similarly, patients with asthma or rheumatoid arthritis who wrote about traumatic experiences showed better health improvements than controls up to four months after the experiment, and participants who spoke about traumatic experiences into a tape recorder were found to have a better immune functioning compared to those who had spoken about trivial topics (Booth, Petrie, & Pennebaker, 1997; Christensen et al., 1996; Smyth, Stone, Hurewitz, & Kaell, 1999). Indeed, the lack of disclosure of emotions and thoughts has been related to deleterious health consequences. Eisenberg and colleagues (2003), using LIWC to investigate the language of HIV-sieropositive women in coping interviews, found that a higher use of emotional expressions and of inhibitory words was significantly associated with lower immunologic responses, after controlling for potentially confounding variables such as demographic factors, race, income and level of education. The increased presence of inhibitory words in articles about human-made traumatic events might then reflect more insidious effects of these disasters, as proposed by Baum's theoretical distinctions (Baum, 1987).

Wikipedia articles about human-made disasters were also characterized by a higher presence of words reflecting social processes. This, along with previous research showing an increase in social sharing and social interactions after traumatizing events (Mehl & Pennebaker, 2003; Pennebaker & Harber, 1993; Rimé et al., 1998), may suggest more insidious and profound effects of disasters caused by

human agencies, which may lead to a language expressing more references to other people and family members. However, to this end it is difficult to draw decisive conclusions, given that in case of man-made disasters the causes are inherently human, and therefore the related articles may contain more references to a killer's social relations or a bomber's family.

Nevertheless, it is relevant to consider that Wikipedia is intended to be an online encyclopedia, and to express a neutral perspective on events, without biases or emotional content ("Wikipedia:Neutral point of view", 2001). Still, the fact that such peculiarities in the use of language emerged from articles about natural and man-made disasters are symptomatic of different psychological processes underlying users' collaboration.

5.4 Discussion

In this chapter the language used in articles related to natural and human-made disasters was analyzed, showing different word usage patterns. The language employed to describe particular events in Wikipedia reflect users' thoughts, emotions and sensemaking activities, and can tell us about the psychological processes at the basis of users' collaboration to articles' editing. In particular, articles about traumatic events caused by human agencies were characterized by an emotional language expressing anxiety and anger, and by a greater presence of words referring to cognitive activity and social processes, with respect to articles about natural disasters. This suggests that even in Wikipedia, where articles are supposed to be written without biases and from a neutral point of view, the collective representation of different types of traumatic events reveals diverse psychological processes. Understanding how different characteristics of traumatic events, such as natural or human-made disasters, are related to particular psychological processes can provide a basis for the prediction of typical social responses, and short and long term effects on individuals and communities (Baum, 1987).

Adopting the theoretical characterization of natural and technological disasters proposed by Baum (1987), the analysis on Wikipedia articles showed that man-made traumatic events may cause more insidious effects. In particular, articles about human-made and natural traumatic events seem to be characterized by specific patterns of emotional language, which in the frame of a dimensional approach to emotions (Lang, Greenwald, Bradley, & Hamm, 1993; Russell, 1980) have a similar negative valence, but opposite levels of arousal (*Figure 5.2*). Although it seems reasonable to find expressions of anger in association to any loss, human-made disasters may engender more anxiety and focused anger because in this case there has been an unexpected loss of control, there is someone to blame, the accident could have been prevented and the victims could have been saved.
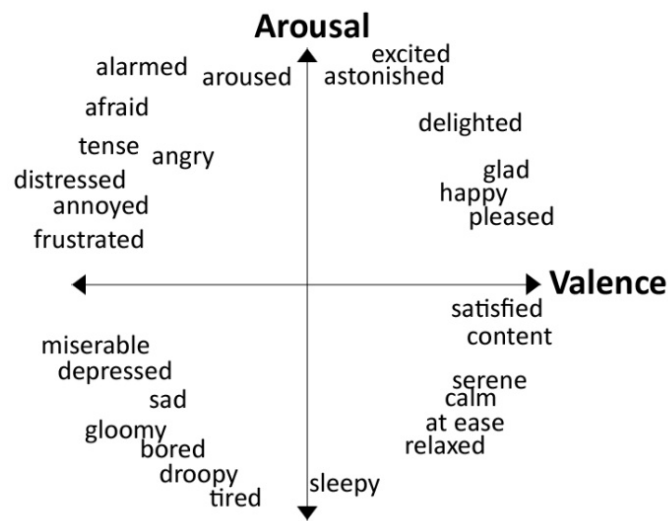
Figure 5.2 In a dimensional model of affect emotions are organized in a two-dimensional space according to their valence and degree of arousal. Adapted from "A circumplex model of affect", by J. A. Russel, 1980, *Journal of Personality and Social Psychology*, *39*(6), p. 1161-1178. Adapted with permission.

For similar reasons, although the need to comprehend the causes and to find a convincing explanation is understandably present in the aftermath of every traumatizing experience (Reser, 2007), it may be even stronger where human agencies are responsible. In articles related to man-made traumatic events, the language associated to cognitive activity such as insight (e.g., "realize", "understand", "think", "consider") and exclusive (e.g., "but", "without", "except", "exclude") words was significantly higher than in those about natural disasters, reflecting increased critical thinking about the reasons of the accidents and cognitive complexity in the construction of coherent narratives. At the same time, the higher presence of inhibition words in articles about accidents where human agencies were responsible supports theoretical suggestions according to which man-made traumatizing events may pose more severe, insidious and long-term effects on the psychological health of people (Baum, 1987; Pennebaker, 1997a). In past research, inhibitory processes were found to be related to poorer health and immunological functioning (Booth et al., 1997; Christensen et al., 1996; McGuire & Kiecolt-Glaser, 2000; Pennebaker, 1993, 1997b; Pennebaker et al., 1989; Pennebaker & Beall, 1986; Petrie et al., 1998; Smyth et al., 1999). In the case of traumatic experiences, the psychological inhibition of thoughts and feelings does not allow to completely process and assimilate the trauma (Lutgendorf & Antoni, 1999).

The fact that distinctive language patterns related to affective and cognitive processes were found in article pages about different traumatic experiences should also be taken into consideration. Indeed, Wikipedia is an online encyclopedia and maintains strict policies regulating article editing. To this regard, one of the most important is the Neutral Point of View policy ("Wikipedia:Neutral point of view", 2001), stating that articles should be written adopting a neutral perspective on events, avoiding biases and emotionally charged content. It is interesting to note that, despite this policy, statistically significant differences in the language expressing affective,

cognitive and social processes can be detected in articles about different traumatic events.

A slightly different explanation could be advanced with regard to the higher presence of words expressing social processes in articles related to man-made disasters. On the one hand, along with previous research findings showing an increase in social sharing and interactions after emotional upheavals, this result suggested that in case of man-made disasters the traumatizing experience may be more insidious, resulting in a language expressing more references to other people and family members. On the other hand, however, such conclusion may be deceiving, given that articles about these kinds of events may simply contain more references to the human agencies at the origin of the disasters.

Indeed, as Pury intelligibly showed (see Back et al., 2010; Back, Küfner, & Egloff, 2011; Pury, 2011), automated text analysis tools like LIWC should be carefully employed when applied to large datasets, especially if they contain automatically generated messages, in order to prevent confounds and hazardous conclusions. Notwithstanding, the massive availability of digital data offers an unprecedented opportunity, if exploited with caution, to access people's thoughts and feelings, allowing scholars to empirically investigate social theories on trauma in ways that would have been difficult just few years ago.

In particular, Wikipedia seems particularly appropriated for the study of the psychological processes underlying the formation of shared narratives about different kinds of traumatic events from two perspectives. On the one hand, the immediate availability of large amounts of data and the fact that the Mediawiki open source web platform powering Wikipedia records any change made by any user to any article or talk page, allow researchers to study the immediate aftermath of single traumatic events. In this way, scholars can assess unobtrusively the presence of different coping strategies, the extent to which various psychological processes are represented in articles and talk pages in the immediate aftermath of an upheaval, and how they may vary in time. On the other hand, the study of how collective memories of traumatic events are formed in Wikipedia, and eventually represented in shared narratives can provide a further high-level perspective on the collective remembrance of emotional upheavals. Collective memories remain alive for years and generations and can influence not only the remembrance of the past, but also attitudes toward the current society (Pennebaker, Paez, et al., 1997). Especially in events of trauma, the study of collective memories can provide insights on current desires, needs and identities of the society, and help to understand how the past can affect the present (Sturken, 1997).

This chapter, analyzing the language used in the English Wikipedia's articles about traumatic events, showed that the shared representation of man-made and natural disasters reveals different kinds of underlying psychological processes. The results suggested that these traumatic events may engender diverse psychological reactions in individuals and communities, possibly requiring different recovery strategy.

The use of automated content analysis techniques, especially if extended to different online environments, could be useful to determine unobtrusively the extent and the nature of distress engendered by collective emotional upheavals and the psychological processes involved in the immediate aftermath and in the construction of the collective representations of these events, as explained in Chapter 6.

Chapter 6

Linguistic markers of emergent memories in Wikipedia talk pages

6.1 Psychological effects and longitudinal analysis of traumatic events

As already discussed in previous chapters, after a traumatic event a common response of the affected populations is to talk about it (Arthur, 2008; Micalizzi, 2012; Pennebaker & Banasik, 1997; Recumber, 2012). Pennebaker and Harber (1993), studying the degree of social sharing of the San Francisco Bay Area earthquake and the Persian Gulf War, found a remarkably high frequency of social sharing and ruminating during the first two weeks after the earthquake or the onset of the war. According to the authors, this early period of increased social sharing laid down the basis for the formation of common experiences and collective memories about the two events.

The social sharing of personal thoughts and emotions serves several psychological purposes. Talking about a traumatizing experience can help people to receive affective and social support, maintaining the cohesion and strengthening the emotional ties among the members of a community (Arthur, 2008; Frijda, 1997; Micalizzi, 2012; Recumber, 2012; Rimé & Christophe, 1997; Wang, 2008). Since talking is a form of rehearsal, the social disclosure of experiences and related feelings can help people to better organize and assimilate events in their minds (Horowitz, 1976). The social sharing following emotional upheavals can also support affected people attempting to understand and to make sense of the trauma (Rimé, 1995).

Past research investigating the social and psychological disturbances after traumatic experiences has documented a generalized increase in the negative mood of people in the immediate aftermath of the events (Koss & Kilpatrick, 2001; Stroebe et al., 2001). As reported in the previous chapter, past studies found persistent anxiety and nervousness (Adler, 1943), emotional problems like depression and personality changes (Green, 1980; Titchener & Kapp, 1976), and also hostility up to several months after traumatic events (Gleser et al., 1981; Gleser et al., 1978; Pennebaker & Harber, 1993).

Also, cognitive activity was found to increase after emotional upheavals, as people engage in rationalization and critical thinking processes to find comprehensive explanations of the events (Davis & Nolen-Hoeksema, 2001). In the case of an unexpected or untimely death, when people feel the urge to find a meaning in what happened, there may be the need for a more philosophical understanding of the loss (Davis & Nolen-Hoeksema, 2001), while in another situations such as technology failures individuals may also seek for causal

explanations (Baum et al., 1983). Analyzing the language used by Mayor Rudolph Giuliani in the course of 35 press conferences between 1993, when he was elected, and 2001, after the September 11 attacks, Pennebaker and Lay (2002) found significant changes in his language style following specific events. In particular, the authors showed remarkable variations in his language after a professional and personal crisis in the late 2000, when in the span of few days he withdrew from the senate race against Hilary Clinton, separated from his wife and was diagnosed with cancer, and after the World Trade Center attacks in 2001. After both of these emotional upheavals he adopted a simpler linguistic style, using less words longer than six letters, but at the same time increased his complexity of thinking. In particular, in the aftermaths of these personal and collective traumatic experiences, he used more causal (e.g., "because", "depend", "effect") and insight and self-reflective words (such as "think", "assume", "interpret", "realize"). Pennebaker and colleagues (Pennebaker, Mayne, et al., 1997; Pennebaker et al., 2003) found that an increased use of language expressing cognitive activity, such as causation words, predicted better health and recovery from traumatic experiences. Studying the role of language as a predictor of health improvement after a traumatic experience, Pennebaker, Mayne, et al. (1997) showed that higher use of self-reflective insightful (e.g., "think", "assume", "interpret", "realize") and causal words (e.g., "because", "depend", "effect") was associated to better physical health, measured as the number of visits to the physician and the rate of physical symptoms reported.

Furthermore, traumatic events have found to be followed by modifications in social interactions, seeking of social support and an increased social orientation (Mehl & Pennebaker, 2003; Pyszczynski et al., 2003; Rimé et al., 1998; Rimé et al., 1992). Using an electronic recorder device, Mehl and Pennebaker (2003) studied students' conversations after the September 11 attacks, comparing them with baseline data previously collected. Although they did not find an overall increase in the amount of social interactions, they found a gradual shift from group conversations toward dyadic interactions, and this change was mildly correlated to better psychological adjustment, measured as the amount of avoidance symptoms, the degree of thoughts and dreams related to the attacks. The authors proposed that one-to-one interactions may have fostered psychological coping processes, providing an intimate context in which students could reinterpret the events and renegotiate their opinions and beliefs. According to Rimé and colleagues (1992), personal and collective traumatic events elicit mental rumination, consisting in frequent thoughts and memories related to the emotional experiences, and these reminiscences are usually shared with others in frequent social interactions. According to the authors, the social sharing of emotions play an important role in consolidating the memories for important events, contributing to the processing of these memories, and enhance social integration and interpersonal relationships (Rimé et al., 1998).

Taking into account past research of the psychological consequences of traumatic events, we hypothesize that an increase in the negative mood, in cognitive processing and in social sharing would be reflected in the language used by people in the aftermath of emotional upheavals. Thanks to the widespread accessibility of digital texts in online environments, it is now possible to study the language patterns expressing people's psychological processes after traumatizing events. The unobtrusive longitudinal analysis of words usage in the aftermath of these events allows scholars to connect individuals' language to specific emotional, cognitive and social processes in the early formation of collective memories (Back et al., 2010;

Cohn et al., 2004; Keegan, 2011; Keegan et al., 2011; Pennebaker & Banasik, 1997; Pentzold, 2009; Tausczik & Pennebaker, 2010).

Cohn and colleagues (2004) used LIWC (Pennebaker et al., 2001) to analyze blog posts during two months prior to and after the September 11, 2001 attacks, and found signs of psychological changes in the language used by bloggers. Specifically, these changes consisted mainly in an increase of words associated to negative emotions, cognitive processing, social engagement, and psychological distancing from the event during the first days following the attacks.

Back and colleagues (Back et al., 2010; Back et al., 2011) employed LIWC to analyze the use of emotional words in messages sent to text pagers in US after the September 11, 2001 attacks, and found an increase in language expressing anger in the first hours after the crashes.

Other recent research focused on the representation of breaking news and traumatic events in Wikipedia. Keegan (2011) selected 127 Wikipedia articles listed under "List of accidents and incidents involving commercial aircraft", classifying them as breaking articles, started within 48 hours of the crash, and non-breaking articles, written more than two days after the accident. In order to understand how editors may embody their psychological states in the articles, he processed every revision with LIWC and analyzed the variations in the use of words with emotional valence, of causal, insightful, tentative and discrepant language, and in the use of narrative immediacy (present vs past tense) as articles stabilized over time. The author found an increase in the emotional valence of articles, a decrease in the use of causal language, and a non significant decrease in the narrative immediacy, showing that while breaking articles tended to become more positive over time, non-breaking articles became progressively more negative, and had an accelerating tendency to contain more past-tense that present-tense language over time.

In order to understand editors' attention and participation dynamics, Keegan et al. (2011) analyzed the patterns of activity on the English Wikipedia related to the 2011 Tōhoku earthquake and tsunami. Considering the category "Tōhoku earthquake and tsunami", the author found 62 articles created before the 11th of March 2011 and 22 articles created after the quake, and showed the impressive collaborative effort in Wikipedia's coverage of the events related to the earthquake, with 3,792 revisions by 1,140 editors to the pre-existing articles, and 11,709 revisions by 2,439 editors to the newer articles. Keegan highlighted an intense attention to the earthquake during the immediate aftermath, followed by a decline and a dispersion of users' interest among other related articles, arguing that editors' response to breaking news in Wikipedia articles followed emergent collaboration practices. In a following study, extending the analysis of collaboration practices to breaking news articles, Keegan and colleagues (Keegan et al., 2011) proposed that the co-authoring activities related to these articles follows novel coordination dynamics which are not exhibited by articles about other historical past events, showing unique collaboration patterns. According to the authors, articles about breaking news show specific collaboration structures, with an initial pattern of highly centralized and clustered interactions among users, which tend to decrease as time goes by and the salience of the events fades, and a well-connected structure of interactions and collaboration among editors of different related articles.

This chapter focuses on two specific traumatic events, the 7 July 2005 London bombings and the 2011-2012 Egyptian revolution, highlighting the temporal evolution of the psychological processes which can be derived from users' linguistic

patterns in the related talk pages in the English Wikipedia. Although both can be considered breaking news articles, they are different in several aspects. The first was a series of co-ordinated terrorist suicide attacks hitting the London transport system during the morning rush hours, started at 8:50 when the first three bombs were detonated on underground trains, and finished about an hour later with a final explosion on a double-decker bus in Tavistock Square. The Egyptian revolution begun with a series of pro-democracy uprisings on 25 January 2011, and was part of a wider revolutionary wave of collective action and demonstrations which lasted several months, involving the Arab world and started in December 2010, with the self-immolation of Mohamed Bouazizi in the town of Sidi Bouzid, Tunisia. These different traumatic events received a substantial attention from the media and attracted intense collaborative efforts on the related articles and talk pages in Wikipedia. The aim of this chapter is to show how the different psychological processes examined earlier evolved over time, analyzing the language expressing affective, cognitive and social processes in the related talk pages (see *Table 4.1*).

6.2 The 7 July 2005 London bombings

On the morning of 7 July 2005, four Islamist terrorists detonated four bombs attacking the London transport system. The first three explosions occurred shortly after 08:50 on three different underground trains, followed by a fourth bomb around an hour later on a bus not far from King's Cross, in Tavistock Square. Fifty-two people were killed in the attacks, including the bombers, and hundreds were injured in what was one of the deadliest attacks in London.

Just few minutes after the first bomb explosion the Wikipedia article on "7 July 2005 London bombings" was initiated, precisely at 09:18. As reported in Chapter 2, just in the course of that first day the article was edited 2581 times. The associated talk page was created few minutes later, at 09:59, and received 626 contributions in its first day.

Brown and Hoskins (2010) analyzed the cultural and mediatic response to the London bombings starting from the work of Goodwin and Gaines (2009), who explored the factors structuring the way terrorist events are appraised by people and their possible consequence for future behavior. They focused on the collective memory of the London bombings and on its relevance for individuals' personal, social, and political concerns, adopting Bartlett's (1932) definition of "schemata" as a set of organized past reactions, which are combined with new information and thus processed to react to the present. Analyzing a corpus of recordings of media coverage about the terrorist attacks, a series of commemorative events, interviews and focus groups with involved people, and digital material such as photographs created by witnesses of the events, the authors proposed that three main schemata were introduced for the interpretation of the bombings. The first was the "Blitz Spirit", which referred to the German bombings of London and other UK cities in 1940-1941. Short after the 2005 bombings, the "Blitz Spirit" was invoked referring to the London citizens strength and willingness to carry on with daily life, as it happened in 1940 when the shops remained open despite the bombings. However, this schemata didn't work in 2005 as many shops closed and the city life froze. The second schemata proposed as an interpretative framework for the terrorist attacks was referred to by the authors as "London Stands United", which recalled ethnicity

and community cohesion issues felt by many Londoners, and the fact that London citizens helped each other during the emergency, proven by digital images and photographs of the immediate aftermath of the attacks. A third schemata referred to the memorial built in 2009 in Hyde park, consisting of 52 stainless steel pillars representing the victims of the bombings. The memorial in Hyde park followed the abstraction tradition typical of the commemorative architecture of the previous decades, such as the Vietnam Veterans Memorial in Washington, DC, and proposed an interpretation of the bombings interweaving individual and collective aspects. Indeed, the memorial highlighted the victims' individuality and at the same time their belonging to a collective of people dead in the same circumstances.

Focusing on collective memory building processes of the London terrorist attacks, Pentzold's (2009) analysis moved to the digital environment considering the communicative and collaboration practices of the English Wikipedia users in the creation of the related article. Using argumentation analysis (Toulmin, 2003), which allowed to highlight conflicting arguments, the author coded by hand the 984 revisions made to the talk page about the London bombings during the two days immediately after the attacks. Besides editorial and technical remarks typical of Wikipedia activity, the major topics were related to the interpretation of the London bombings as "terrorism", to the comparisons to other attacks such as the Madrid bombings or the Lockerbie bombing, described in Wikipedia in the article on "Pan Am Flight 103", to eyewitness accounts of people working or traveling in London, and to messages expressing condolences, compassion and confusion about the bombings.

In this section, drawing from previous research findings on the psychological consequences of traumatic events (Cohn et al., 2004; Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Pyszczynski et al., 2003; Rimé et al., 1998; Rimé et al., 1992; Stroebe et al., 2001), we hypothesize a higher presence of language expressing affective, cognitive and social processes (see *Table 4.1* for the details about the related LIWC categories) during the immediate aftermath of the bombings in the English Wikipedia talk page associated to the attacks. In particular, a higher initial presence of language referring to negative emotions would confirm the strong emotional impact of the bombings on the article editors, despite Wikipedia promoting a detached and neutral perspective ("Wikipedia:Neutral point of view", 2001) and despite, according to Wikipedia guidelines and policies, talk pages should not be used to advance personal views or to discuss about the subjects of related articles ("Wikipedia:What Wikipedia is not", 2001; "Wikipedia:Talk page guidelines", 2005). An increased presence of words indicating cognitive activity is also expected in the first days after the bombings, pointing at users' critical thinking processes in order to find comprehensive explanations about what was happening, who the perpetrators were and their reasons. Finally, a higher presence of language referring to social processes during the immediate aftermath of the bombings would confirm the increased initial social orientation and the presence of social sharing of thoughts, emotions and feelings which has been found in the past works described earlier.

6.2.1 Method

Single edits to the talk page of the article about the London bombings during the immediate aftermath after the attacks, from 7 July to 12 July, were automatically

extracted and processed with LIWC. Only the first 6 days were considered because, despite an initial substantial participation, the talk page showed a gradual and constant decrease in users participation, typical of breaking news articles (Keegan, 2011). As already mentioned in Chapter 3, talk pages are usually less edited with respect to article pages, serving users' need to discuss the articles' content ("Wikipedia:What Wikipedia is not", 2001; "Wikipedia:Talk page guidelines", 2005). Indeed, while the article about the bombings received 2581, 777, 276, 189, 169, and 158 edits during the first six days, the related talk page received respectively 626, 363, 57, 69, 32, and 29 edits, as shown by *Figure 6.1.*
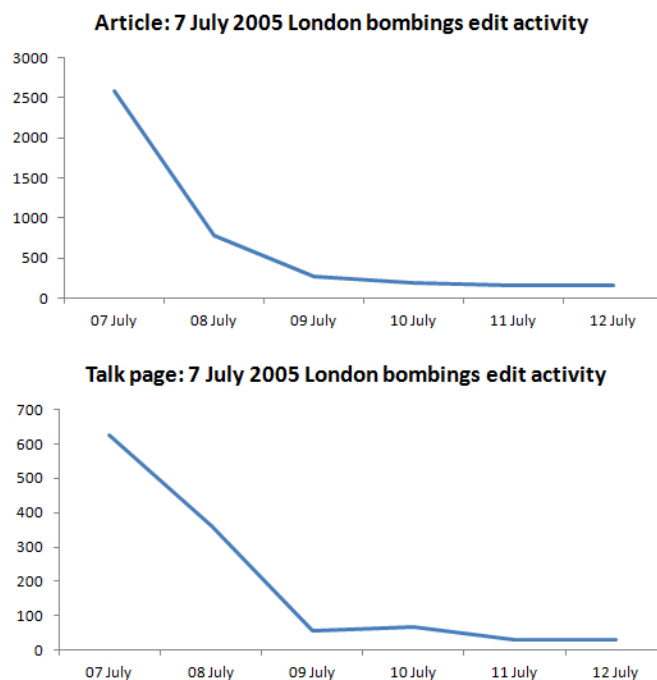

Figure 6.1 Number of edits per day to the article and talk page of the "7 July 2005 London bombings".

For each edit only content newly added by users was considered, with the purpose of avoiding pieces of preexisting text automatically included in the revisions. Adopting the same procedure as of Back and colleagues (2010), in order to reduce the data variability and to obtain pure measures for each of the psychological categories of interest reported in *Table 4.1*, the corresponding scores were computed as percentages of words in relation to the total number of words recognized in the dictionary.

Automatically analyzing textual data often requires some preprocessing routines (Back et al., 2010; Back et al., 2011; Pury, 2011). In the case of Wikipedia revisions, data preprocessing was needed because of the presence of technical codes inserted by users and other automatically added content, such as templates, signatures and timestamps.

Users' signatures were removed since they do not typically convey relevant information about the content of the edits, but nevertheless they may contain words which could wrongly lead to confounds in the edit interpretation. Signatures were automatically removed exploiting the common Wikipedia syntax. For example, the preprocessing script removed pieces of codes such as [[User:Rich Farmbrough|Rich

Farmbrough]], displayed on Wikipedia as a link to the personal page of "Rich Farmbrough", the second most active user on the article about the London bombings. In fact, in the reported piece of syntax LIWC would recognize the two words "Rich", assigning them to the category of positive emotions. However, since the syntax used for signing posts and for other templates can be added both exploiting the icons provided in the edit toolbar or by hand, in this latter case the code added by users is more difficult to automatically detect, and might have remained even after the preprocessing.

Since in some cases the automatic exclusion of signatures could not perform well, for instance when users signed their posts by hand and mistakenly included syntax errors in their signatures, minimal changes were operated in the LIWC dictionary. These modifications consisted in removing the names of months, which are automatically added in the Wikipedia timestamp of each edit to identify the date and hour of a revision, and which were irrelevant for the purposes of the study. Furthermore, the case of "May" could have led to possible confounds, since it was assigned by LIWC dictionary to the categories of tentative language and cognitive processing, and thus would have probably produced misleadingly high percentages of words related to cognitive processes in the month of May. For this reason, the word "May" with the capital "M" was removed, maintaining on the contrary the word "may" with the small "m".

Moreover, shorter revisions, which were to a large extent pieces of signatures or templates slipped through the automatic deletion procedure, contained few words recognized in the dictionary. This would have caused a high variability in the percentages computed over recognized words. For example, considering an edit for which only one word is recognized (e.g., "good"), this edit would produce a score of 100 % in the related category (in this case, positive emotions). In order to reduce this spurious variability caused by shorter edits, only revisions with at least six words recognized in the dictionary were considered for the analysis. This threshold was found to be a fairly good compromise allowing to exclude irrelevant pieces of content, without at the same time removing too many short, meaningful edits. Considering edits with at least six words recognized in the dictionary reduced the dataset to daily samples of 425, 251, 39, 47, 21, and 19 revisions. A few examples of revisions with less than six words recognized in the dictionary are reported in *Table 6.1*.

Revision	Words recognized in the dictionary
- Serialized	0
\n~~ fourthed Apsmif101	1
(Sound Off) 04:40, 10 July	2
("see also") ("external links")	3
Casuality Information / Sources ==\n(Renamed from "at L ")	4

Table 6.1 Examples of edits containing less than six words recognized in LIWC dictionary and therefore excluded from the analysis.

The need for preprocessing and cleaning procedures just explained provides the opportunity to highlight the complexity of the revisions extraction processes and the subsequent analysis of edits containing automated text with word count approaches, as intelligibly discussed by Pury (Back et al., 2011; Pury, 2011). Indeed, despite their effectiveness for studying large datasets, automated text analysis tools should be carefully employed, especially when the analyzed corpus contains automatically generated messages which, when repeated at a high frequency, could transform a secondary measurement error into serious confounds in the data interpretation. These misleading confounds could be prevented by examining in detail samples of text reporting high percentages for specific categories, and looking for false positives. In the case of the present study on the temporal evolution of linguistic patterns in the "7 July 2005 London bombings" and the "2011-2012 Egyptian revolution" talk pages, this recursive investigation on the performance of automatic revision extraction and LIWC analysis allowed to elaborate the preprocessing procedures explained above, aimed at excluding automatically generated content.

After we obtained the texts of single revisions, LIWC scores were calculated for the psychological categories of interest already reported in *Table 4.1*, and data were subsequently aggregated per day, for the first six days. Since the data distributions, despite the arcsine transformation (explained in Chapter 4) did not meet the assumption of normality, tested with the Kolmogorov-Smirnov test, the original scores were analyzed with the Kruskal-Wallis test, a rank-based method for analyzing the distributions of more than two independent groups with unequal sample sizes (Motulsky, 1999; Wilcox, 2009). Similarly to the analysis of variance, the statistical significance of the Kruskal-Wallis test implies a significant difference between at least two of the groups analyzed (Motulsky, 1999; Sheskin, 2004), and therefore Games-Howell post-hoc tests were applied when appropriate for further explanation of specific differences. The Games-Howell test was found to be one of the most robust methods for unequal sample sizes and when the assumption of normality is violated (Hilton & Armstrong, 2006).

6.2.2 Results

The graphs in *Figure 6.2* report the results with regard to the categories of affective processes, confirming the hypothesis of a decreasing trend over time. The Kruskal-Wallis tests were significant for the overarching category of affective processes ($\chi^2_{(5)}$=20.27, p=.001), for negative emotions ($\chi^2_{(5)}$=19.39, p=.002), anxiety (e.g., "worried", "afraid", "apprehensive"; $\chi^2_{(5)}$=12.64, p=.027) and anger (e.g., "kill", "aggression", "destroy"; $\chi^2_{(5)}$=14.28, p=.014), while it did not attain statistical significance for positive emotions (e.g., "love", "party", "pleasant"; $\chi^2_{(5)}$=8.15, p=.002) and sadness (e.g., "sad", "cry", "depression"; $\chi^2_{(5)}$=1.73, p=.885). For affective processes, Games-Howell post-hoc tests confirmed a statistically significant difference between the scores of the last (12 July) and the first three days (7, 8, 9 July; respectively with p<.001, p<.001, and p=.001) and between the scores of the fourth (10 July) and the first couple of days (7 and 8 July; respectively with p=.009 and p=.012). A similar pattern of statistically significant differences was found for negative emotions, which differed between the last (12 July) and the first three days (7, 8, 9 July; respectively with p<.001, p<.001, and p=.001) and between the scores of the fourth (10 July) and the first couple of days (7 and 8 July; respectively with

p=.007 and p=.001). The scores related to anxiety were significantly different between the first day after the attacks (7 July) and the fourth, fifth and sixth days (10, 11, 12 July; respectively with p=.002, p=.047, and p<.001), and between the second day (8 July) and the third, fourth, fifth and sixth (9, 10, 11, and 12 July; respectively with p=.020, p=.001, p=.009, and p<.001). Finally, Games-Howell post-hoc comparisons showed statistically significant differences for anger between the sixth (12 July) and the first, second and third days (7, 8, 9 July), respectively with p=.001, p<.001, and p=.015).
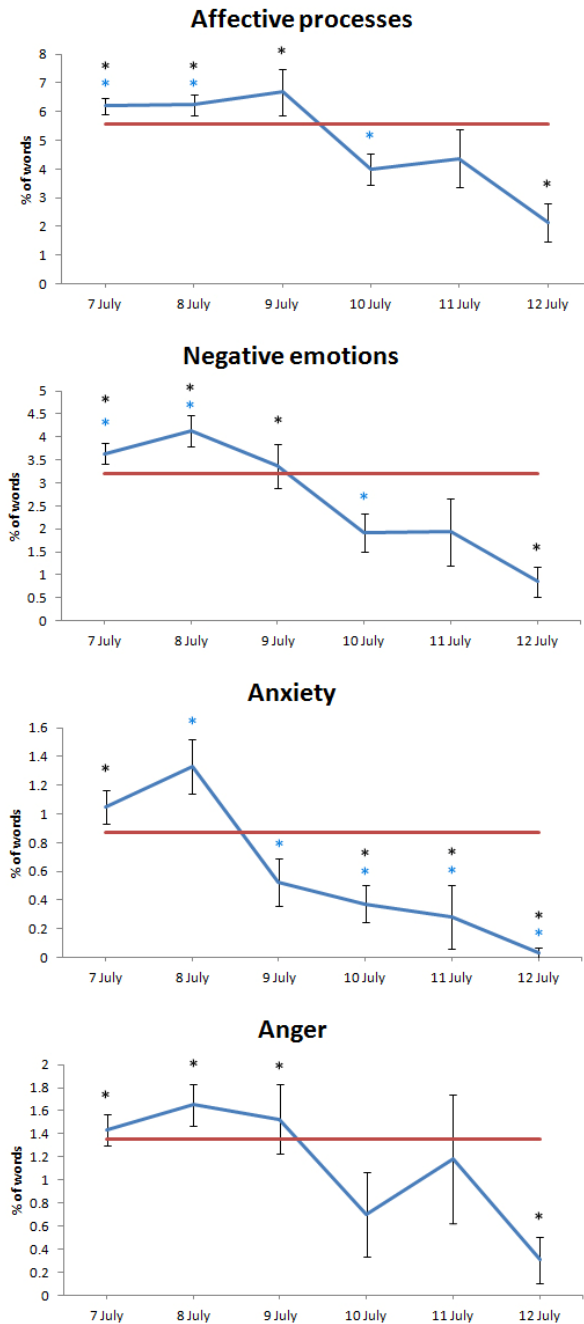

Figure 6.2 Time trend of the language related to affective processes, negative emotions, anxiety and anger in the talk page about the "7 July 2005 London bombings", during the first six days after the attacks. The mean scores of words

75

expressing affective processes, negative emotions, anxiety and anger, represented with the red lines, were computed considering the entire revision history of the talk page until the data collection, at the end of March 2011. The temporal trends of categories which did not reach statistical significance are not reported.

These results confirmed the hypotheses of an increased presence of language expressing affective processes, and in particular anxiety and anger, during the immediate aftermath of the London bombings, validating previous literature findings (Gleser et al., 1981; Green, 1980; Koss & Kilpatrick, 2001; Pennebaker & Harber, 1993; Stroebe et al., 2001; Titchener & Kapp, 1976). The presence of significantly higher amounts of affective content during the first days reflected the intense emotional impact of the London bombings among Wikipedia editors who, besides posting about condolences and mourning for the dead, expressed their nervousness, anxiety and fear also making sharp comments and sometimes losing their temper, as showed in the following pieces of discussions[4]:

> **Oh fuck**
> Sorry. Just had to say that. -- 7 July 2005 10:12 (UTC)
>
> **Early comments**
> Leaving all things aside, I need to say this is shocking, just shocking. Also, I and my family feel eerily de-sensitised to the idea of a terrorist-like act. Harro5 July 7, 2005 09:59 (UTC)
>
> […] is this helpful??
> I'm safe. Its chaos over here and I'm sh***ing myself. The bombs are really close, oxford street has been closed down and there are still bombs planted everywhere waiting to go off. If ed and di can get out of the central london bit, tell them to just stay away and away from any form of transport. The hospital is chaos and we have been put on high alert. Its nuts, but we are safe, apart from the fact I have just s**t myself, and I serioulsy mean it. We aren't going to be able to get home tongith because we've been told there are bombs everywhere. I'm scared, but I'm sure it will all be okay.
> Source unknown...
>
> **Condolences**
> I am so sorry to hear about this terrible and barbaric act of terrorism. My thoughts are with the British people. --mav 7 July 2005 13:11 (UTC)
>
> Thanks for that. Mine too. My thoughts with those who have suffered in this nightmare -- Grace Note
>
> Condolences to the families and everyone affected. Best wishes to the injured and emergency staff to prevent more loss of life. --213.54.220.28 7 July 2005 15:27 (UTC)

---

[4] The full text of the comments can be found at
http://en.wikipedia.org/w/index.php?title=Talk%3A7_July_2005_London_bombings&diff=18319202&oldid=18319137,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_1#Early_comments,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_3#Condolences,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_2#Protected.3F.21,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_4#Terrorism.2C_part_2.

[…] I offer my condolences to those who are dead, dying, or hurt, and to the families of the before mentioned. --Admiral Roo July 7, 2005 16:05 (UTC)

[…] Deepest condolences from Slovenia. Terrible not only for London, but for whole Europe. I stand with you. --Eleassar my talk 7 July 2005 17:47 (UTC)

My condolences go out to the British people on this tragic incident.Kiwidude July 7, 2005 21:05 (UTC)

[…] Condolences from Slovakia. I am so sorry to hear about this barbaric act of terrorism. Our thoughts are with the British people! --Ondrejk 7 July 2005 18:28 (UTC)

This is truly a horrible event. My condolences go out to the victims and their families. --Ixfd64 2005 July 7 19:08 (UTC)

Count me in as one who's thoughts go out to all Londoners, and especially those that were touched directly by these horrible events. AlvinMGO 7 July 2005 19:37 (UTC)

[…] Condolences from Morocco for this crazy act of terror to famillies and everyone affected. I am so sorry for the victims, injured and wounded londoners. My thought for all thouse experiencing this, it's so penfull to be victim of absurd war, --Mjiyed 7 July 2005 23:01 (UTC)

Protected?!
Who the hell just protected the page? People are trying to update! --Nidonocu 7 July 2005 13:17 (UTC)

Terrorism, part 2
Once again this dispute arises. People have also argued about whether the the 9/11 attacks should be called "terrorism." I don't think this should even be an issue. A fairly simple definition of terrorism is that it's deliberate violence unlawfully perpetrated against civilian targets for a political purpose. […] Mr. Billion 8 July 2005 00:22 (UTC)

[…] That's my point--there is little use for the term "terrorism" on an NPOV encyclopedia. It's not denotation, it's connotation. Please READ MY FUCKING ARGUMENTS. — Phil Welch 8 July 2005 00:25 (UTC)

There isn't any reason to lose your temper. Moncrief July 8, 2005 00:29 (UTC)

I was speaking in the tradition of "read the fucking manual", but I think there's PLENTY reason to lose my temper. You people are shouting me down without reading, understanding, or responding to my arguments. — Phil Welch 8 July 2005 00:31 (UTC)

Quite surprisingly, the results related to cognitive processes did not confirm our expectations, showing lower percentages of insight words during the first days, with a significant increase in the last day of the analysis for insightful words (*Figure 6.3*). Indeed, Kruskall-Wallis tests were significant for insight (e.g., "think", "assume", "interpret"; $\chi^2_{(5)}$=20.28, p=.001), causal (e.g., "because", "depend", "elicit"; $\chi^2_{(5)}$=13.43, p=.020) and words expressing certainty (e.g., "always", "absolutely", "clear"; $\chi^2_{(5)}$=18.50, p=.002), while it was not significant for discrepancy (e.g.,

"should", "could", "if"), tentative (e.g., "maybe", "apparently", "suppose"), inhibition (e.g., "block", "abstain", "avoid"), inclusive (e.g., "and", "add", "along") and exclusive (e.g., "but", "either", "without") language (respectively, $\chi^2_{(5)}$=7.64, p=.177; $\chi^2_{(5)}$=3.17, p=.674; $\chi^2_{(5)}$=4.11, p=.534; $\chi^2_{(5)}$=8.43, p=.135; $\chi^2_{(5)}$=8.38, p=.136). Although the respective Kruskall-Wallis tests were significant, Games-Howell post-hoc comparisons did not attain statistical significance for causal words and words expressing certainty. This could suggest a possible lack of statistical power to perform the post-hoc comparisons. However, Games-Howell post-hoc tests showed a significant difference between the scores of insightful words in the first (7 July) and the last day considered in the study (12 July; p=.040).
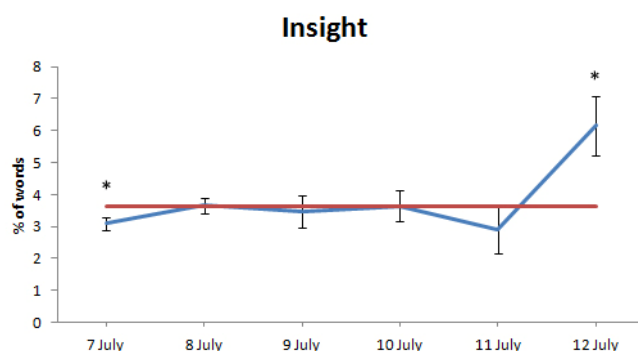

Figure 6.3 Time trend of the language related to insight processes in the talk page about the "7 July 2005 London bombings", during the first six days after the attacks. The mean score, represented with the red line, was computed taking into consideration the entire revision history of the talk page until the data collection, at the end of March 2011. The temporal trends of categories which did not reach statistical significance are not reported.

These results, showing a significant increase of insightful language during the sixth day after the attacks, suggested that, when emotional language was at its height in the immediate aftermath of the attacks, cognitive activity remained rather constantly around the mean, increasing only a few days later, when on the other side expressions of negative emotions decreased. It is possible that after a few days, when heated debates and sharp comments started to tone down, Wikipedia users had the possibility to concentrate on article editing, such as sourcing and references activities, and on the related discussions on the talk page, as well as on understanding in detail the succession of events, as showed by the following comments, which reported a high presence of words related to cognitive processes and insight[5]:

> **Similar incidents**
> The list of "Similar incidents" is growing; it's in danger of becoming a
> category in it own right, I suggest it is drastically trimmed (perhaps with a

---

[5] The full text of the comments can be found at
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_6#Similar_incidents,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_6#References_please,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_6#identified_Casualties
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_6#How_to_include_these_references.3F,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_6#Timing_of_trains.

link to a suitable umbrella article or category). Andy Mabbett 17:10, 12 July 2005 (UTC)

References please
Can we please have a reference for the police conference about the investigations into four men seen on CCTV at Kings Cross ("Investigation" section). If this is properly referenced, then the earlier speculation about a single bomber (based on tube geography) in the "Bombings" section can be removed. Carcharoth

identified Casualties

[…] Yes, Gladys Wundowa has disappeared from news coverage, so presumably there was a mistake. She shouldn't be on the list (or at least there should be a note by her name) until more is known. --Dtcdthingy 13:19, 12 July 2005 (UTC)

How to include these references?
Under "Bombings" section I've added detail about the other trains at Edgware Road at the time of the bombings. This is mainly based on the images available here (TrackerNet images from London Transport):[11] These additions are supported by the BBC reference I gave in the main article […], plus this reference to the Hammersmith and City Line train at Edgware Road at the time of the bombing […]
I've included the BBC link and the link to the Transport for London news centre image gallery, where people can see the sequences showing the movements of the trains, but can't think of an easy way to reference the H+C line comment. Any way to do this? Carcharoth 21:37, 12 July 2005 (UTC)

Timing of trains
In the "Attacks on the Underground" section, is the speculation about a lone bomber justified? It starts out as speculation, and then phrases like "the bomber", and "the person who planted the other bombs" start to be used, which sound less speculative and start to assume that the speculation is correct, which doesn't seem justified yet.
The wording "roughly nine minutes east by tube from Kings Cross" and "roughly eight minutes west by tube from Kings Cross" is also a bit strange. London Underground have a system called TrackerNet that allows them to tell PRECISELY when the Circle Line trains arrived at and left Kings Cross […]. So IF London Underground release the information, there is no need for "roughly"; it can be said that the trains left Kings Cross "x" minutes ago. I thought one of the original press releases from London Underground said both Circle Line trains left Kings Cross 8 minutes before the bombings, but that seems to have disappeared now. Can anyone find another source?
[…] Carcharoth 08:08, 12 July 2005 (UTC)

Regarding to social processes (see *Figure 6.4*), only the scores of category related to the friendship domain (e.g., "buddy", "friend", "mate") were found to be statistically different across days ($\chi^2_{(5)}$=15.25, p=.009), while the overarching category of social processes and the subcategories related to family (e.g., "daughter", "brother", "dad") and humans (e.g., "adult", "children", "girl") did not attain statistical significance (respectively, $\chi^2_{(5)}$=8.88, p=.114; $\chi^2_{(5)}$=6.64, p=.249; $\chi^2_{(5)}$=6.99, p=.222). With regard to the friend category Games-Howell post-hoc comparisons showed a significant difference between the first (7 July) and the third and sixth days (9 and 12 July), respectively with p=.025 and p=.022. The score represented by the

peak in *Figure 6.4* and corresponding to the fifth day after the attacks (11 July) was not significantly higher than others, probably due to the high variability. Indeed, an examination of the revisions scores frequencies showed a right skewed distribution (N=21; M=0.21; SD= 0.54), with several revisions containing no words related to friends, and 3 revisions respectively with 1.67, 1.61 and 1.2 % of words referring to the friendship domain.
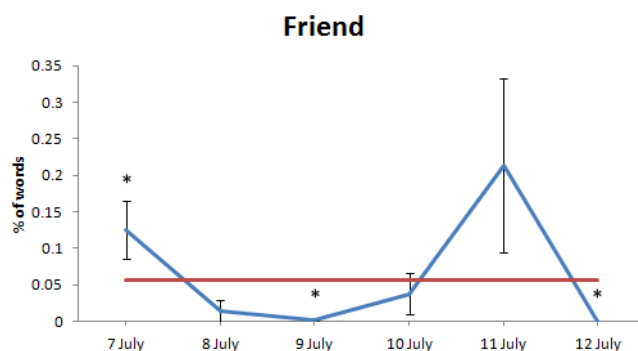

Figure 6.4 Time trend of the language related to the friendship domain in the talk page about the "7 July 2005 London bombings", during the first six days after the attacks. The mean score, represented with the red line, was computed considering the entire revision history of the talk page until the data collection, at the end of March 2011.The temporal trends of categories which did not reach statistical significance are not reported.

These results partially validate the initial increase in social orientation of people affected by a traumatizing experience, during its immediate aftermath. Indeed, the higher presence of words related to the friendship domain in the day of the attacks seems to suggest the preoccupation of Wikipedia users for their friends and relatives, as confirmed by the following comments, which were categorized by LIWC's dictionary as showing high percentages of words related to friends[6]:

> **contacting friends / families etc.**
> Does anyone know of a coordinated attempt to get word of people who are safe to their friends and relatives - something similar to what was used during the tsunami etc. ? -- User:Jdowland
>
> […] Those 2 dead could be anybody's friend or family member they haven't heard from. Let alone those who are injured and the unconfirmed dead. The system used during the tsunami was supposed to be an enormous success and re-assured a great number of people. The same principle could be applied to this and other events. -- Jdowland
>
> **Airports?**
> Has there been any reports on whether airports are operational? I have a friend who is supposed to be flying in, in a few days.

---

[6] The full text of the comments can be found at
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_1#contacting_friends_.2F_families_etc.,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_1#Airports.3F,
http://en.wikipedia.org/wiki/Talk:7_July_2005_London_bombings/Archive_4#Hotline.

> […] The radio just mentioned that, I think they are staying open. Your friend should have no trouble.-- Kizor 7 July 2005 11:35 (UTC)
>
> **Hotline**
> The hotline numbers etc have disappeared. I've no great problem with this, but perhaps others do? Evercat 7 July 2005 17:23 (UTC)
>
> It kind of helps to have the numbers there, as people are trying to find out information about relatives, friends, etc. and it saves a lot of work looking up the numbers. Kaiser Matias 18:47 7 July 2005 (UTC)

Summarizing, the results confirmed an initial increased presence of language related to affective processes, and especially to anxiety and anger, suggesting that during the immediate aftermath of the attack the emotional load was at his height, resulting in mourning comments for the deceased, but also triggering nervousness and anxiety which can be inferred by the heated debates reported above. During this initial period characterized by emotional language, the words expressing cognitive activity, and in particular insight remained rather steadily around the mean, to increase after a few days, on the 12th of July. On the one hand, this results was unexpected, considering past studies finding an increase in cognitive processes after traumatic events (Cohn et al., 2004; Pennebaker & Lay, 2002). In particular, Cohn and colleagues (Cohn et al., 2004), who used LIWC to analyze the linguistic temporal patterns of blogposts after the September 11 attacks, found a significant increase of words expressing cognitive activity during the 11th and 12th of September, followed by a rapid decrease within the span of a week. On the contrary, in the Wikipedia talk page related to the London bombings, the increase of insightful language was found six days after the attacks, when the emotional language had dropped below the mean. This could suggest that, at least in the case of the talk page on the London bombings, in the immediate aftermath users were still concerned with the emotional impact of the traumatizing experience and therefore there was less space for critical thinking. After a few days, when the mourning messages remembering the victims had been posted and the heated debates had calmed down, it is possible that Wikipedia users could concentrate better on article editing activities, and on finding detailed explanations about the events. Finally, the results related to social processes partially confirmed past literature findings about the increased social orientation after traumatic events, although in the case of the London bombings users' initial concerns seemed to be directed at finding information about friends and relatives, and helping others to contact their families.

6.3 The 2011-2012 Egyptian revolution

At the end of 2010 and for the subsequent months, a series of pro-democracy rebellions and demonstrations started to spread across North Africa. The first protests occurred in Tunisia following Mohamed Bouazizi's self-immolation in the town of Sidi Bouzid on 17 December 2010, and quickly started to propagate in other countries such as Egypt, Algeria, Libya, Lebanon, Jordan, Palestine, Iran and Bahrain. Soon, the Middle East was shocked by a wave of violent demonstrations, riots and strikes, rapidly resulting in the overthrow of two heads of state, the presidents of Tunisia Zine El Abidine Ben Ali and of Egypt Hosni Mubarak.

A key role in the context of North African uprisings has been played by the use of social media and social networking sites (Beaumont, 2011; Chebib & Sohail, 2011; Khamis & Vaughn, 2011). New Internet-based information systems allowed novel ways for accessing, producing and circulating information, leading to the reconfiguration of communication systems on a global scale, and to new forms of individual expression, based on decentralized and horizontal communication (Castells, 2001). Technologies based on Web 2.0 environments allowed new forms of participation in the social and political life, fostering freedom of expression and the circulation of protests on a large scale (Burns & Eltham, 2009; Chebib & Sohail, 2011; Grossman, 2009; Khamis & Vaughn, 2011; Morozov, 2009). Twitter, for instance, played a crucial role in the spreading of protests in Iran after the presidential elections in 2009, when Mahmoud Ahmadinejad won against the opposition candidate Mir-Hossein Mousavi (Burns & Eltham, 2009; Ferron, in press; Grossman, 2009; Morozov, 2009). The death of Khaled Said, killed by the Egyptian police in the summer of 2010, would have maybe passed unnoticed if Facebook, Youtube and Twitter had not been so popular. At October 2012, the Facebook page "We Are All Khaled Said" counted around 250 thousands "likes", and Khaled Said became a symbol of the Egyptian revolution (Owais, 2011; "We Are All Khaled Said," 2010).

Chebib and Sohail (2011) discussed the role of social media in the Egyptian revolution. According to the authors, specific characteristics allowed social media to play a key role in the revolution. For instance, they are easily accessible by everybody without licenses or fees, they are user-friendly and easy to figure out even without instructions. They allow ubiquitous access because people can use mobile devices to take pictures and videos and upload them on the Internet, where their visibility can be amplified by mass media. Social networking sites such as YouTube provide easy ways to embed videos on blogs, newspaper and television websites, or other social networking websites. For these reasons, social media have contributed to democratize power and social influence, pulling down communication barriers and favoring instantaneous information sharing, updated in real time. Fostering the spreading of uncensored and unedited information, they opened the way for participatory journalism, allowing ideas to spread rapidly in densely connected networks and providing people with new opportunities of participating in the protests, making them more visible (Chebib & Sohail, 2011; Khamis & Vaughn, 2011). According to Khamis and Vaughn (2011) social media also helped promoting civic engagement, because their virtual platforms allowed activists to discuss, exchange ideas, plan and organize protests. And furthermore, they allowed citizens to document the protests and to provide evidence of the government brutality, disseminating ideas, pictures and videos to the world outside Egypt (Khamis & Vaughn, 2011).

Abel and colleagues (Abel, Gao, Houben, & Tao, 2011) employed user modeling strategies to analyze users' activity on Twitter, in order to understand their interests. They monitored more than 20,000 users over a period of more than 4 months, starting in November 2010, and collected more than 30 million tweets, analyzing how users' interests in specific topics changed over time, focusing especially on the Egyptian revolution. The authors analyzed the temporal evolution of the interest in the Egyptian revolution modeling topics on Twitter through a set of weighted concepts, with concept defining arbitrary entities represented through named entities or hashtags (such as #jan25 or #tahrir), and weights reflecting the importance of the

concept for the topic. Some entities like "Cairo" and "Mubarak" were popular for the topic over a long period of time, meaning that Twitter users kept on referring to these entities in their tweets about the Egyptian revolution. Specifically, the occurrence frequencies of these entities reached their peak on the 28th of January, three days after the first protests, and then begun to decrease slowly over the next couple of weeks. In contrast, some other entities like "SMS", "Omar Suleiman" and "Vodafone" showed a different pattern of high spikes, indicating they were relevant for the topic only for short periods of time. For instance, the entity "SMS" showed an occurrence peak on the 28th of January and referred to the shutdown of the Internet and of short messaging service by the government. Similarly, "Omar Suleiman" became popular on the 30th of January, after he was appointed as vice-president on the 29th of January, causing a further wave of protests. The entity "Vodafone", which registered a spike of occurrence on the 3rd of February, likely referred to the news that the company announced about the Egyptian government hijacking their network. Considering a subset of users for which they extracted at least 20 tweets in total and 10 tweets during the period of the revolution, the authors found that 70% of users showed interest in the Egyptian revolution, but only less than 150 started tweeting about the revolution from its very beginning on the 25th of January. On the contrary, about 450 users joined the discussion on the topic between the 28th (Friday of Rage) and 29th of January, after the Egyptian authorities shut down the Internet on the 26th. The authors also analyzed for how long users remained interested in the topic, distinguishing between short-term and long-term adopters. Short-term adopters published a fairly high amount of messages in a short time period, usually less than a week, while long-term adopters tweeted about the revolution for more than a month. The authors suggested the presence of a correlation between the time when users started to be involved in the Egyptian revolution and the duration of their interest, observing that early adopters tended to overlap stronger with long-term than with short-term adopters, although they did not test this correlations with statistical analyses. Moreover, they observed that short-term adopters seemed to be more strongly influenced by the news media than long-term adopters: indeed, while short-term adopters showed a peak of interest after the "Battle of the Camel" on the 2nd of February, which was widely described by news reports and heavily discussed in the social media, long-term adopters showed more frequent peaks of interest, even in days which were not characterized by epic events.

As the protests and demonstrations of the Egyptian revolution spread across the streets with the participation of a multitude of people who used social media like Facebook or Twitter to organize the uprisings, a less visible share of activity happened also on the online encyclopedia Wikipedia. Thousands of people contributed to the related Wikipedia pages, which were promptly created as the uprisings started to spread, trying to make sense and represent them as they unfolded over time.

Ferron and Massa (2011a) provided evidence of this intense activity on different language versions of Wikipedia pages, and analyzing excerpts of discussions in the article and users talk pages, in which editors negotiated different perspectives on the events, proposed to interpret the phenomenon as a process of collective memory building. The exploratory analysis of the discussion network on the talk page about the Egyptian revolution allowed to identify key editors playing different social roles and thus contributing in specific ways to the editing activities and to the discussion.

Heather Ford (2012) studied users' participation and coordination activities on the English Wikipedia article and the talk page about the Egyptian revolution from the creation of the article (25th of January) to the 12th of February, after Mubarak resigned and the article name changed from "2011 Egyptian protests" to "2011 Egyptian Revolution". Ford showed that during the first 18 days of protests, editors from Egypt, United States, Middle East, Western Europe and elsewhere worked together to build a shared and accurate representation of the events as they were unfolding. According to the Pew's Research Center's Project for Excellence in Journalism (PEJ), the coverage of the Egyptian uprising from 31 January to 6 February was the biggest in the past four years, filling 56% of the newshole (Jurkowitz, 2011). It was also the biggest overseas story in such few days since PEJ began its News Coverage System in 2007, and the fourth biggest story in general. There are several reasons which can explain this utmost interest in the coverage of the first days of the protests. According to (Ford, 2012), one reason was the number of people and journalists documenting events with videos and images, such as the "Battle of the Camel" on the 2nd of February, when Mubarak supporters on camels attacked the protesters in Cairo. Another was the uncertainty of the outcome of the protests, about the possible resignation of Mubarak and about who would have ruled the country after him. Also the role of media drew further attention to the uprising, making the story even more compelling, with journalists being threatened, harassed, and in some cases detained. According to Ford, there were two additional reasons for the Wikipedia community to be so active in building an accurate version of the events. First, the unprecedented shut down of the Internet services by the Egyptian government during the initial days of the protests led Wikipedia users to actively work to protect Internet's freedom from the government's attacks. Secondly, some of the editors were directly involved, living in Egypt or having friend and relatives there. For many of these users, participating in the construction of a collective representation of the protests was a way of being involved in these important historical events. The user who initially created the article, who was also one of the most active editors, wrote on the article talk page[7]:

> I completely understand your concern. Some of my friends have hurt during the protests so I might too invested in the article. I have been trying to be as neutral as I much as I can I am still a human. That's why I asked for more editors to come and help to make sure I am not leaning on way or another. It helps that I blow most of anger on twitter and facebook when I get or read bad news. But I think I owe the people who are protesting and those who have passed away to tell their story from a NPOV. I am not nor can I be in Egypt due my university so that my way of standing up to Mubarak and tell him I shall not be silenced and I shall tell the world what happened. I would be happy not to take the lead. That why I asked Lihaas among others to add certain parts that I know I can not do. like the wael part due to my personal relationship with him. I hope you would join us in editing tho. If you maintain a NPOV, trust me, other editors will let you know (me included). P.S Always happy to meet another fellow Deist-- The Egyptian Liberal (talk) 09:17, 10 February 2011 (UTC)

---

[7] The full text of the comments in the related discussion can be found at
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_6.

In the comment, "The Egyptian Liberal" was probably referring to his relationship with Wael Ghonim, the creator of the Facebook page "We Are All Khaled Said", who was reported missing and secretly detained by the Egyptian government after his arrival in Cairo (Caulfield, 2011; Owais, 2011; Whitlock & Fadel, 2011). Furthermore, the need to be involved in the story retelling and in the construction of the collective memory of the uprising expressed by "The Egyptian Liberal" in the comment above motivated many other users and recalls the directive function of memory proposed by Wang (2008).

In this work, considering the English Wikipedia talk page associated to the Egyptian revolution, and based on previous studies on the psychological outcomes of traumatic events (Cohn et al., 2004; Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Pyszczynski et al., 2003; Rimé et al., 1998; Rimé et al., 1992; Stroebe et al., 2001), we hypothesize a higher presence of language related to affective, cognitive and social processes (see *Table 4.1*) during particular days of the revolution. Unlike the London bombings, which were a series of blasts temporally confined in the span of few hours, the Egyptian revolution was a series of protests and uprisings started on 25 January 2011, but subsequently unfolded in the course of several months. Here we focus on the first days of protests, hypothesizing an increase in the presence of language expressing affective, cognitive and social processes in the most heated, and therefore most important, days of the revolution.

Considering the patterns of users edit activity to the talk page (*Figure 6.5*), we identified the most heated days as the 3rd and the 11th of February. *Figure 6.5* shows the number of edits per day to the English Wikipedia article and talk page of the "2011-2012 Egyptian revolution".
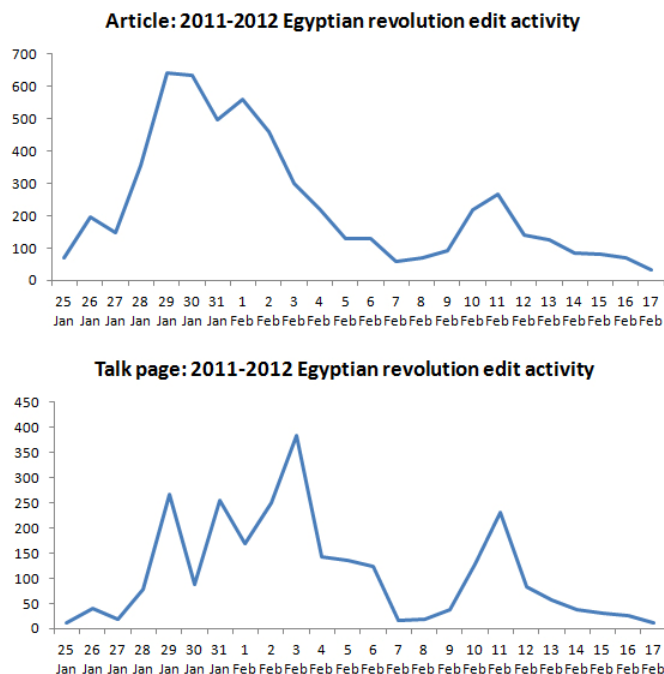

Figure 6.5 Number of edits per day to the English Wikipedia article and talk page of the "2011-2012 Egyptian revolution".

While the article edit activity reached its highest peak during the initial days of the revolution, a second increase in the number of edits can be identified on the 11th of

February, the day Hosni Mubarak resigned. Expectedly, the associated talk page showed in general a lower amount of activity. Though, it followed a similar pattern, with the higher share of participation during the first days of protest culminating in a peak around the 3rd of February, and a subsequent increase on the 11th of February.

A review of the main events which characterized the first days of the Egyptian revolution can further highlight the most heated moments of the initial uprisings, as showed in the timeline represented in *Figure 6.6* (AlJazeera, 2011; BBC, 2011; Blight, Pulham, & Torpey, 2012; Maher & Eskandar, 2012; von Hein, 2011; Wolman, 2012).

The timeline shows an escalation of violence, started with the initial protests in Tahrir Square on 25 January, and culminated on the 2nd and 3rd of February in the "Battle of the Camel", when violent clashes burst between pro-Mubarak supporters on camel and horses and the protesters. According to the timeline of events, the other most significant day of the revolution was on the 11th of February, when finally Hosni Mubarak stepped down, handing over the power to the army and producing widespread wild celebrations across the country.

The patterns of edit activity to the talk page (*Figure 6.5*), and the timeline reported by international newspapers and broadcasting media (*Figure 6.6*), suggest that two of the most important moments in the revolution were the 2nd-3rd and the 11th of February, two days considered crucial for the series of pro-democracy protests which shook Egypt in the early 2011.

Taking into account past research on the psychological outcomes of traumatic events, we hypothesize that the emotional charge of the "Battle of the Camel" and of the "Defiance Friday" would be reflected in specific language patterns in the Wikipedia talk page of the "2011-2012 Egyptian Revolution". In particular, an increased presence of affective language around the 2nd-3rd and the 11th of February would confirm the high emotional impact of the related events on users working to reach a shared representation of the revolution. An increased presence of words expressing cognitive activity would suggest critical thinking processes in order to find out what was happening and the possible outcomes after Mubarak's resignation. Finally, an increase in the language referring to social processes around the 2nd-3rd and the 11th of February would reveal the high social orientation expected in the most heated days of the revolution.

Egyptian revolution timeline

25 Jan	"Day of Rage". A series of protests start in Tahrir Square, Cairo, and rapidly spread across the country. Police use violence, tear gas and water cannon to disperse the protesters.
26 Jan	AlJazeera reports the first protester and police officer killed in Cairo. Clashes are reported in Suez.
27 Jan	ElBaradei joins the protests, which continues across the country. Facebook, Twitter and other telephone and Internet services are disrupted.
28 Jan	"Friday of Rage". While the disruption of Internet services continues, Egypt prepares for new protests after Friday prayers.
29 Jan	Mubarak refuses to step down in a speech to the nation, and the demonstrations continue in Tahrir Square despite the police attempts to disperse the protesters.
30 Jan	Demonstrations continue and protests are reported in Alexandria.
31 Jan	Mubarak still refuses to step down and thousands of protesters gather in Tahrir square.
1 Feb	"March of Millions". Mubarak promises to reform the constitution, in particular the article preventing independent candidates to run for office. People flood in Cairo to take part in the "March of Million".
2 Feb	"Battle of the Camel". While Internet services are partially restored, pro-Mubarak supporters on camels and horses enter Tahrir Square and attack protesters, who fight back to keep the square.
3 Feb	Violent clashes between supporters and opponents of Mubarak continue to spread in Tahrir Square. The confrontations resulting from the "Battle of the Camel" have been reported by BBC News as the worse since the beginning of the revolution.
4 Feb	"Day of Departure". Protesters gather in Tahrir Square asking Mubarak to step down.
5 Feb	Thousands of people remain in Tahrir square, while the leadership of the National Democratic Party resigns, including Gamal Mubarak, son of Hosni Mubarak.
6 Feb	While the protests continue in Tahrir square, the banks re-open for a few hours in the attempt to get the city start returning to normal.
7 Feb	While thousands of people continue camping in Tahrir square and schools remain closed, the activist Wael Ghonim is released after a few days of detention.
8 Feb	As the government announces a plan for a peaceful transfer of power, Tahrir Square holds the biggest crowd of protesters.
9 Feb	Protesters continue to join the streets and Tahrir Square, and massive strikes roll throughout Egypt.
10 Feb	The newly appointed Culture Minister, Gaber Asfour, quits. When Mubarak states he will remain in power until September, the protesters react with fury weaving their shoes in the air and asking the army to join the demonstrations.
11 Feb	"Defiance Friday", or "Friday of Departure". After the demonstrators marched across the streets and under the presidential palace in angry protests, Hosni Mubarak resigns and hands over the power to the army.

Figure 6.6 Compendiary timeline of the most important events characterizing the Egyptian revolution. The events were reconstructed according to articles published in international newspapers and media (AlJazeera, 2011; BBC, 2011; Blight et al., 2012; Maher & Eskandar, 2012; von Hein, 2011; Wolman, 2012).

6.3.1 Method

The method adopted for the automatic extraction of revision and for analysis with LIWC was the same used for the study on the talk page about the "7 July 2005 London bombings" (see Section 6.2.1). Single edits to the talk page about the "2011-2012 Egyptian revolution" from 25 January to 17 February 2011 were automatically extracted and subsequently processed with LIWC. For each revision, only newly added content was extracted in order to avoid preexisting text. Signature templates were removed and only edits containing at least six words recognized in the dictionary were considered, with the purpose of reducing the variability of the scores and excluding pieces of automatically generated content. Single revisions were analyzed using LIWC dictionary from which the name of the months had been removed, in order to exclude irrelevant content which had slipped through the signatures deletion. The scores corresponding to the main psychological categories of interest (see *Table 4.1*) were computed as percentages of words in relation to the total number of words recognized in the dictionary (Back et al., 2010). Since the number of revisions varied substantially across days, ranging from few tens to several hundreds of edits as showed in *Figure 6.5*, the data were aggregated in 12 groups of two subsequent days (respectively with N=27, N=69, N=237, N=244, N=400, N=152, N=98, N=31, N=228, N=75, N=37, N=28), in order to reach an adequate number for each group allowing for statistical analysis. Aside from this, the rest of the procedure remained unchanged.

Since, despite the arcsine transformation already employed in Chapter 4, the data distributions did not approximate normality, assessed with the Kolmogorov-Smirnov test, the original scores were analyzed with the Kruskal-Wallis test (Motulsky, 1999; Wilcox, 2009). Games-Howell post-hoc tests were applied when the Kruskal-Wallis test revealed significant differences across groups, to account for specific differences (Hilton & Armstrong, 2006).

6.3.2 Results

The graphs in *Figure 6.7* report the temporal trend and the results related to the categories of affective processes, anger and sadness in the talk page about the Egyptian revolution, from 25 January to 17 February 2011. The Kruskal-Wallis tests were significant for the categories of negative emotions ($\chi^2_{(11)}$=20.10, p=.044), anxiety (e.g., "worried", "afraid", "apprehensive"; $\chi^2_{(11)}$=32.22, p=.001), anger (e.g., "kill", "aggression", "destroy"; $\chi^2_{(11)}$=23.74, p=.014) and sadness (e.g., "sad", "cry", "depression"; $\chi^2_{(11)}$=28.19, p=.003). Games-Howell post-hoc test did not reveal significant differences across groups for the words related to anxiety. However, they showed statistically significant differences between the scores of 16-17 February and 31 January - 01 February, 2-3 February and 4-5 February for negative emotions (respectively with p=.014, p<.001, and p=.026); between 16-17 February and 29-30 January, 31 January - 01 February, 2-3 February, 4-5 February and 10-11 February for words expressing anger (respectively with p<.001, p<.001, p<.001, p=.010, and p<.001); between 2-3 February and 25-26 January, 27-28 January and 16-17 February (respectively with p=.001, p=.002, and p=.001), and between 10-11 February and 25-26 January, 27-28 January and 16-17 February for words related to sadness (respectively with p=.002, p=.004, and p=.002).
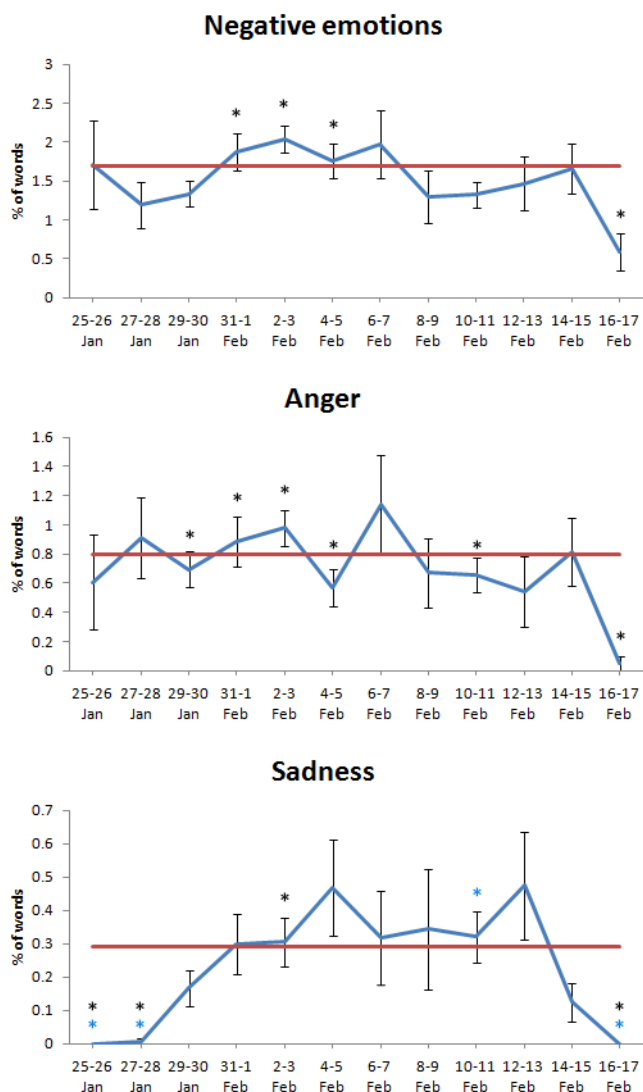
Figure 6.7 Time trend of the language related to negative emotions, anger and sadness in the talk page about the "2011-2012 Egyptian revolution", from 25 January to 17 February 2011. The mean scores of words expressing negative emotions, anger and sadness, represented with the red lines, were computed considering the entire revision history of the talk page until the data collection, at the end of August 2011. The temporal trends of categories which did not reach statistical significance are not reported.

In particular, the amount of words expressing negative emotions and specifically anger was found to be above the average and significantly higher during the days surrounding the "Battle of the Camel", with respect to the last day considered in the study. Language expressing sadness was significantly higher during the "Battle of the Camel" and on the day Hosni Mubarak resigned, in comparison to the first and the last days. A further consideration of the revisions posted during the days

characterized by a higher presence of negative emotions in the language used by Wikipedians can highlight the topics under discussion[8]:

Attacks on the media
The Committee to Protect Journalists has collated a list of attacks on the media since the protests began: Mubarak intensifies press attacks with assaults, detentions. Chesdovi (talk) 18:55, 3 February 2011 (UTC)

Military not neutral, has been kidnapping and torturing protesters
[…] Here's some more sources on this topic.
"Torture by Egypt army reported" - Ottawa Citizen
"Military Accused of Torture, Abuse" - Wall Street Journal
"Egypt's military accused of torturing anti-government protesters" – Sify
Is that good enough? SilverserenC 03:53, 11 February 2011 (UTC)

OR table of "Confirmed death toll" of 382 is back, once again[9]
It's totally incorrect, absolutely false, it's a lie, it isn't and can't be backed by any reliable source […] --94.246.150.68 (talk) 14:20, 3 February 2011 (UTC)
I don't know who is doing it, even the table's author admitted it was sloppy math. The other random, unsourced figures for the so-called "Confirmed death toll" besides 382 included 351, 410, 376 and more - all equally simply invented by Wikipedia user(s) attempting to do things clearly not allowed by WP:OR (read it, dammit). --94.246.150.68 (talk) 14:30, 3 February 2011 (UTC)

[…] Get rid of this thing, and keep it out, until you get the confirmed figures by reliable sources. --94.246.150.68 (talk) 14:39, 3 February 2011 (UTC)

[…] No. Original. Research. Allowed. On. Wikipedia. Get it? -- 94.246.150.68 (talk) 14:56, 3 February 2011 (UTC)

[…] It's not a matter of consensus, it's matter of Wikipedia policy: NO ORIGINAL RESEARCH ALLOWED ON WIKIPEDIA, NEVER, ANYWHERE, AT ALL. […] So, again it's a lie. --94.246.150.68 (talk) 15:30, 3 February 2011 (UTC)

---

[8] The full text of the comments in the related discussions can be found at
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#Attacks _on_the_media,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_7#Military _not_neutral.2C_has_been_kidnapping_and_torturing_protesters,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#OR_tab le_of_.22Confirmed_death_toll.22_of_382_is_back.2C_once_again,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_6#Calm_t alking.2FWikipedia:What_Wikipedia_is_not...,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_2#.22A_p oster_of_Khaled_Mohamed_Saeed.22,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_5#Wael_ Ghonim,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_6#Request ed_move.
[9] This discussion, archived at
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#OR_tab le_of_.22Confirmed_death_toll.22_of_382_is_back.2C_once_again, reported a template referring to the Wikipedia guidelines ("Wikipedia:Etiquette," 2001; "Wikipedia:Staying cool when the editing gets hot," 2002), and asking users to keep calm and respect others.

Stop calling people lairs. Geez -- The Egyptian Liberal (talk) 15:33, 3 February 2011 (UTC)

[…] STOP CAT-FIGHTING, YOUR NOT A BUNCH OF RIOTERS IN TIRIR SQUAIR!!!!!!!!!!!!!!!!!!! (Sorry, but it had to be said!)Wipsenade (talk) 15:24, 3 February 2011 (UTC)

Calm talking/Wikipedia:What Wikipedia is not...
Let's calm down a bit and stop flameing off about Egypt, please! 12:08, 4 February 2011 (UTC)

[…] Let's all read the WP:CIVIL page! User:The Egyptian Liberal, User:Lihaas and IP number User:94.246.150.68 should stop cat fighting on Talk:2011 Egyptian protests. It may be best to investigate some of the disruptive IPs and trowaway acounts and lock the page to IPs for a month or two.

[…] This is pretty outrageous where you guys are asking people to calm down while people are being openly killed while defenseless, and many major media are downplaying the killings, or just not covering it hardly at all. […] —Preceding unsigned comment added by 67.169.72.154 (talk) 13:09, 7 February 2011 (UTC)

The role of Al-Ahly ultras in the protests
[…] So people are getting killed, thousands wounded, the economy is collapsing etc etc etc - and you wish to mention that a few Soccer games have been cancelled?Moxy (talk) 07:23, 4 February 2011 (UTC)

"A poster of Khaled Mohamed Saeed"
[…] Khaled Saeed was one of the main catalysts for these protests and his name is a rallying cry for the protesters. SilverserenC 01:07, 2 February 2011 (UTC)

Wael Ghonim
A group of the youth organizers in the liberation Square name Wael Ghonim to speak on their behalf. Wael Ghonim has been missing since Friday of Anger. -- The Egyptian Liberal (talk) 20:06, 2 February 2011 (UTC)

[…] I actually know Wael personally. I hope he is alive, I really do. Its very sad. There is only two options: Dead or arrested. I hope its the later so at least they will release him. sry, its bit hard to talk about this or write about it -- The Egyptian Liberal (talk) 06:33, 3 February 2011 (UTC)

Requested move
I would agree that, at this point, revolution as a title seems to be the only option. I think it's quite clear with Mubarak's resignation that this should be considered a revolution. Since his resignation an hour ago, almost all news sources have been referring to it as a revolution. SilverserenC 17:34, 11 February 2011 (UTC)

The comments posted on different topics reported above suggest that the linguistic analysis detected both users documenting military violence against journalists and other people, such as the posts in "Attacks on the media" and "Military not neutral, has been kidnapping and torturing protesters", but also heated discussions among editors, reported in the subsequent comments. With regard to sadness, it seemed that users were referring to one of the symbols of the revolution,

Khaled Said, who had been killed by the police the previous summer, and to Wael Ghonim, the creator of the Facebook page "We Are All Khaled Said", who at the time of the comments was still detained. However, a further examination of the comments posted on 11 February suggested that the higher presence of language related to sadness on the day of Mubarak's resignation may be due to a deceiving interpretation of LIWC's dictionary. Indeed, the word "resign" and its inflections are assigned by the dictionary to the category of sadness, and since Mubarak's resignation was the topic of several messages posted to the talk page on 11 February, this led to a misleading interpretation. Once again, this emphasize the need for further control routines when implementing automatic text analysis and especially word count approaches, which can be powerful tools for analyzing large amount of data but are not sensitive to context.

The results related to cognitive processes followed an unexpected temporal pattern, but consistent with previous results related to the London bombings (*Figure 6.8*). The Kruskal-Wallis tests revealed significant differences across groups for the categories of words expressing discrepancy (e.g., "would", "should", "could") and certainty (e.g., "always", "absolutely", "clear"; respectively, $\chi^2_{(11)}$=29.40, p=.002 $\chi^2_{(11)}$=21.88, p=.025), but Games-Howell post-hoc comparisons did not show significant differences between any pair of groups for words related to certainty. On the contrary, discrepancy words were found to be significantly higher on 29-30 January, with respect to 2-3 February (p=.044).
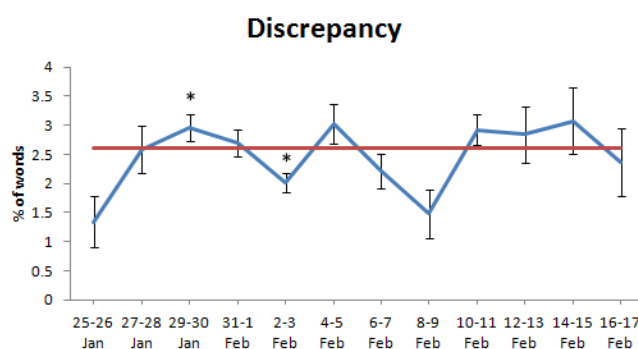

Figure 6.8 Time trend of the language expressing discrepancy in the talk page about the "2011-2012 Egyptian revolution", from 25 January to 17 February 2011. The mean score was computed considering the entire revision history of the talk page until the data collection, at the end of August 2011. The temporal trends of categories which did not reach statistical significance are not reported.

Some comments posted on the 29-30 January on the talk page about the "2011-2012 Egyptian revolution" and characterized by a high presence of discrepancy words can highlight the content of the revisions[10]:

---

[10] The full text of the comments in the related discussions can be found at
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_2#Political _cartoon_NPOV_Issue,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_1#Pendin g_changes_re-enabled,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_1#Venezu ela,

Political cartoon NPOV Issue
I think putting both of them in the Reception section would mitigate any POV problems. SilverserenC 00:42, 29 January 2011 (UTC)

Pending changes re-enabled
Agree that the volume of vandalism is manageable. Pending changes are not needed. If anything, maybe semi-protection but I don't see enough problems for that either. --Aude (talk) 02:05, 29 January 2011 (UTC)

Venezuela
Should the takeover of the Egyptian embassy in Venezuela be added to the article? 65.93.15.80 (talk) 09:24, 29 January 2011 (UTC)

time to call it Revolution?
If and when Mubarak leaves, then it becomes an "Uprising", Not a Revolution. We need have NPOV. If and when it becomes a full-blown revolution, then we will name it a "Revolution" -- The Egyptian Liberal (talk) 08:06, 30 January 2011 (UTC)

This result, contrary to our hypotheses but consistent with previous findings on the talk page of the "7 July 2005 London bombings", may actually confirm the interpretation previously provided. Indeed, also in the case of the "2011-2012 Egyptian revolution", words related to cognitive processes were significantly higher when emotional language was lower. Specifically, discrepancy language (e.g., "would", "should", "could") has been related to linguistic politeness (Sexton & Helmreich, 2003). This could mean that, while the uprisings were at their beginning and the situation seemed to be unstable but still under control, Wikipedia editors worked at the article showing polite manners, trying to select meaningful information managing sources and references. After a few days, however, the situation in Egypt started to deteriorate with violent clashes in the "Battle of the Camel", harassment and attacks to journalists, and at the same time the discussions on Wikipedia became heated and sometimes rude, as showed by previous comments expressing negative emotions and anger.

With regard to social processes, the Kruskal-Wallis tests showed the presence of statistically significant differences across groups for the overarching category of social processes (e.g., "we", "you", boy"; $\chi^2_{(11)}$=24.10, p=.012) and for the subcategory of words referring to human beings (e.g., "adult", "children", "girl"; $\chi^2_{(11)}$=20.06, p=.045). Confirming the hypotheses, Games-Howell post-hoc comparisons revealed significantly higher percentages of words related to social processes during the days corresponding to the "Battle of the Camel", on 2-3 February, particularly with respect to 10-11 February (with p=.011; see *Figure 6.9*). Similarly, confirming our expectations, the language referring to the human beings domain was found to be significantly higher during 31 January – 1 February, 2-3 February, and 4-5 February, with respect to 16-17 February (respectively, with p=.021, p<.001, and p=.021), and an additional significant difference was reported between 2-3 and 10-11 February (p=.022), with language related to humans being higher during the 2-3 February, surrounding the "Battle of the Camel".

http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_3#time_to _call_it_Revolution.3F.
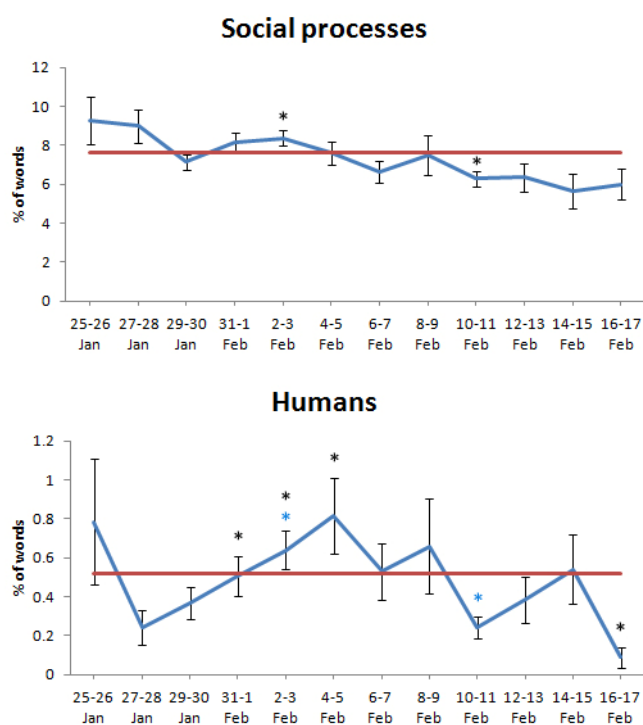
Figure 6.9 Time trend of the language related to social processes and human beings in the talk page about the "2011-2012 Egyptian revolution", from 25 January to 17 February 2011. The mean scores, represented with the red lines, were computed taking into consideration the entire revision history of the talk page until the data collection, at the end of August 2011. The temporal trends of categories which did not reach statistical significance are not reported.

The comments reported below, characterized by a high presence of language related to social processes and human beings and posted by Wikipedia users during the days surrounding the "Battle of the Camel", reveal an increase in social interactions between users, probably associated to the heated debates and arguments mentioned before, but also an increased orientation towards the role of journalists and women in the protests and to the attacks to human rights[11].

> **What to do**
> No, we need to integrate it elsewhere, then delete. (There's little valuable info there anyway.) And try to remember we're working to trim the main article, not expand it now. --94.246.150.68 (talk) 11:25, 3 February 2011 (UTC)
>
> and whose WE? because I know Im not -- The Egyptian Liberal (talk) 11:28, 3 February 2011 (UTC)

---

[11] The full text of the comments in the related discussions can be found at
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#What_to_do,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#The_role_of_women,
http://en.wikipedia.org/wiki/Talk:2011%E2%80%932012_Egyptian_revolution/Archive_4#Number_of_journalists_attacked.3B_human_rights_members_attacked.

OK. "We" except you. --94.246.150.68 (talk) 11:35, 3 February 2011 (UTC)

everyone above disagree with (YOU). so again whose WE? -- The Egyptian Liberal (talk) 13:43, 3 February 2011 (UTC)

And who is "everyone above" that disagree with "(ME)"? "We" - we, the community. Editors of Wikipedia. The issue of article trimming, fighting overlink, timeline sub-article, etc, was raised repeatedly by various editors. --94.246.150.68 (talk) 13:56, 3 February 2011 (UTC)

The role of women
Women protesters number anywhere from 20% to 50% participation, depending on the time of day. I think this is relevant, considering that traditional participation of Egyptian women in public life is less than 10%. Is anyone interested in including a section about women's participation? […]
USchick (talk) 17:24, 4 February 2011 (UTC)

Number of journalists attacked; human rights members attacked
This article from Forbes discusses that 24 reporters were arrested on Thursday and 21 were assaulted. It also explains that at least 30 human rights group members, such as Amnesty International members, were arrested on Thursday. […] SilverserenC 04:09, 4 February 2011 (UTC)

In summary, the hypotheses on an increased presence of language related to affective, cognitive and social processes on the days surrounding the "Battle of the Camel" and Mubarak's resignation were partially confirmed. In particular, no significant results were found on the 11th of February, with the exception of a significantly higher amount of words related to sadness which, however, seemed to be mainly due to a misleading interpretation of expressions referring to the former president's resignation. On the contrary, a significantly higher presence of emotional language and of words related to social processes was found around the 2-3 February, when violent clashes between pro-Mubarak supporters and protesters took place in Tahrir Square. This referred, on the one hand, to Wikipedia users documenting and discussing specific events inherently traumatic, such as attacks to journalists, the violation of human rights and the military accused of torture. On the other hand, however, the increased emotional language and words associated to social processes suggested the escalation of heated discussions among editors, which more than once required the intervention of other users to soothe the exceedingly violent arguments. Interestingly, the pattern showed by discrepancy language did not confirm the hypotheses but was consistent with previous results related to the talk page of the London bombings. Similarly to what happened in the case of the "7 July 2005 London bombings" talk page, discrepancy language, an indicator of cognitive activity and specifically of politeness (Sexton & Helmreich, 2003), was significantly higher when the emotional language was low, at the end of January, few days before the "Battle of the Camel". Considering the comments posted by Wikipedians during that days, this could signify a productive coordination activity among users, who tried to identify and combine important information for the article. As the protests on the streets started to grow violent, however, the communicative interactions became less polite and users started to raise the tones of the discussions.

6.4 Discussion

This chapter analyzed the language used in two talk pages of the English Wikipedia dedicated to the London bombings of 2005 and to the Egyptian revolution started in 2011, specifically focusing on the temporal evolution of particular psychological processes detectable in users' linguistic patterns. Although these events differed in many aspects, both were highly traumatizing. Taking into account past research on the consequences of traumatic events, we hypothesized a higher presence of language associated to affective, cognitive and social processes during the immediate aftermath of the London bombings and during the days characterized by the most violent and significant events for the Egyptian revolution, which according to the patterns of edit activity on the related talk page, and to the timelines reported by traditional mass media, were identified as the 2nd and the 11th of February, corresponding to the "Battle of the Camel" and the resignation of Mubarak.

In general, the hypotheses were partially confirmed, with some unexpected results. In particular, as expected, emotional language was found to be significantly higher in the immediate aftermath of the London bombings and during the days surrounding the "Battle of the Camel" for the talk page about the Egyptian revolution, while in this case no significantly higher presence of affective words were found around the 11th of February. For the talk page about the London bombings, emotional language referred to comments expressing grief and mourning for the deceased, but also anxiety and anger. For the talk page about the Egyptian revolution, words expressing anger were related both to users' comments reporting violent episodes, but also to an escalation of heated debates among users, while language related to sadness referred to Khaled Said and Wael Ghonim, symbols of the revolution. The higher presence of words associated to sadness during the 11th of February seemed to be due to a deceiving interpretation of the word "resign" and its inflections, which referred to Mubarak handing over the power to the army. This misleading results provided the opportunity to recall the main criticism moved to word count approaches, discussed in Chapter 4, about their blindness to context.

One unexpected result, but consistent across the two studies, was related to the different temporal evolution of language related to affective and cognitive processes. Contrary to previous findings showing a parallel increase of words related to emotional and cognitive activity (Cohn et al., 2004), the linguistic analysis of Wikipedia talk pages associated to the London bombings and to the Egyptian revolution suggest a specular evolution which is graphically summarized in *Figure 6.10*. One possible explanation of these contrasting results could be the different digital environments analyzed. Cohn and colleagues (2004) considered journal entries retrieved on the blog hosting social network livejournal.com. These entries consisted in users' personal diaries, which are not typically conveyed in a collaborative environment. When posting on their online diaries, users disclose their thoughts, reflections, feelings and emotions in an individual context. Comments to their posts may be allowed, but fostering interactions among users is not the primary objective of blog hosting digital platforms. On the contrary, Wikipedia talk pages were created specifically to provide users a shared space where they could discuss and coordinate edit activity. Therefore, in this communal environment, social processes involved in human interactions could influence the temporal evolution of the psychological processes identifiable in users' linguistic patterns. In this context, the social dimension may influence the emotional sharing after traumatic events, exacerbating

arguments and inhibiting other psychological processes related to cognitive activity, such as critical thinking and linguistic politeness.
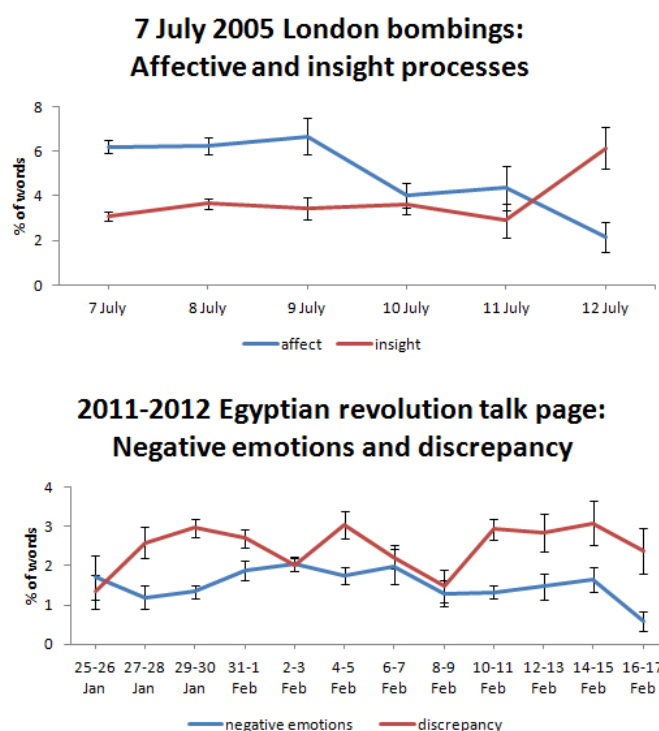

Figure 6.10 Temporal trend of language related to affective and insight processes in the talk page about the "7 July 2005 London bombings", and temporal trend of language related to negative emotions and discrepancy in the talk page about the "2011-2012 Egyptian revolution".

As suggested by the graphs reported in the figure and by the statistically significant differences reported in the results sections, it seems that while the emotions were high, language expressing cognitive activity was lower, and when insightful and discrepancy words increased, the emotional language decreased. Specifically, in the case of London bombings it seemed that, while in the immediate aftermath users were expressing the emotional impact of the attacks by posting condolences and mourning messages and showing nervousness and anxiety, after a few days, when the heated debates had calmed down, they could concentrate on article editing and on finding explanations to the events. Similarly, in the case of the Egyptian revolution, when the protests were still unfolding, Wikipedia editors worked at the article showing the linguistic politeness expressed by a higher presence of discrepancy words (Sexton & Helmreich, 2003), but when the uprisings grew violent, during the days of the "Battle of the Camel", some of them lost their temper and also the language became more frantic and passionate.

Finally, confirming the hypotheses, language related to social processes was found to be significantly higher during the immediate aftermath of the London bombings, suggesting the immediate concern of users for helping people to find information about friends and families. Words referring to social processes and to human beings were also found to be higher in the days surrounding the "Battle of the Camel", in the talk page about the Egyptian revolution, showing both heated

discussions and an increased orientation towards the role of women in the uprisings and the journalists being harassed.

In summary, the results partially confirmed past research findings about increased negative mood, cognitive and social processes after traumatizing events, but revealing specific temporal patterns different from those showed by Cohn and colleagues (2004). A further investigation of specific psychological processes involved in the immediate aftermath of emotional upheavals could allow a better comprehension of the outcomes of traumatic events and of their psychological consequences. In the long run, the immediate reactions to traumatizing experiences will contribute to form the related collective memories, which are able to affect the society and attitudes toward the present (Pennebaker & Harber, 1993; Sturken, 1997). To this end, the use of automated text analysis techniques could be useful, along with the necessary employment of careful control routines, to analyze large amount of data available in digital environments such as Wikipedia, providing researchers with unique opportunities to unobtrusively access people's thoughts in real time.

Chapter 7

Conclusions

The unifying theme of this dissertation has been the investigation of collective memories and the shared representation of traumatic events in the English Wikipedia applying quantitative methods. First, I analyzed edit activity patterns to show the presence of collective memory building activities in Wikipedia. Then, I presented empirical studies on the psychological processes detectable in the representation of different traumatic events, and from users' interactions in the talk pages of two traumatizing experiences, the 2005 London bombings and the 2011 Egyptian revolution.

Through the manuscript, I addressed four research questions, which are discussed below.

Can Wikipedia be interpreted as a global memory place, and does it encompass the typical processes of collective memory building?

The first research question was addressed in Chapter 3. The quantitative analysis of edit activity during anniversaries of traumatic events, supported by the content of users' comments in the talk pages, confirmed the presence of commemoration practices in Wikipedia. In particular, the relative amount of edit activity during anniversaries was a statistically significant predictor in discriminating pages about traumatic events from other pages, showing that Wikipedia articles and talk pages dedicated to emotional upheavals are generally modified to a greater extent during anniversaries. This suggested the presence of dynamic commemoration practices typical of collective memory building, which convey the "lack of closure" of trauma (Sturken, 1997, p. 63) and the need for temporal orientation (Frijda, 1997) through the periodical re-interpretation of past events. Anniversaries of traumatic events in Wikipedia become the occasion to re-negotiate divergent perspectives on the past, and to bring the shared narrative framework into question in order to update it. Anniversaries become also moments for grieving and mourning (E. Zerubavel, 2003), for expressing anger, frustration, and fear. The comments in the talk pages analyzed in Chapter 3 suggested the presence of the socio-affective and memorative functions of collective memory during the anniversaries (Arthur, 2008; Frijda, 1997; Micalizzi, 2012; Recumber, 2012; Wang, 2008), represented by users' verbalization and social sharing of emotions, by their prayers for the victims and their families, and by their demands for justice and respect. The directive function of memory, discussed by Frijda (1997) and Wang (2008), was particularly manifest in the edits to the talk page of the Virginia Tech massacre article during its first anniversary, which

showed users' awareness of the educating power of Wikipedia in the prevention of similar tragedies in the future. Furthermore, the collaborative dimension concurs to distinguish Wikipedia from other digital archives and to define it as a global memory place (Pentzold, 2009). Contrary to Facebook groups, blogs, Twitter or online memory banks such as the September 11 digital archive, which are fragmentary collections of individual contributions (Arthur, 2008; Micalizzi, 2012; Recumber, 2012), in Wikipedia different interpretations of the past are selected and combined into coherent narratives through a dynamic and collaborative process which is typical of collective memory building (Jedlowski, 2001; Pentzold, 2009; Sturken, 1997). The commemoration patterns in the edit activity during anniversaries, together with the other functions of collective memory identifiable in users' comments in the talk pages about traumatic events, confirmed Wikipedia as a global memory place, where collective memories are built though discussions and discursive interactions among editors.

Are word count approaches effective in detecting differences in the Wikipedia representations of traumatic and non traumatic events, and in the temporal focus of old and recent emotional upheavals?

The second research question was addressed in Chapter 4, which discussed the use of word count approaches to detect specific psychological processes in articles about traumatic and non traumatic events, and a different temporal focus in articles about recent and old traumatic events. LIWC (Pennebaker et al., 2001) was employed for the textual analysis of Wikipedia pages for three main reasons. First, to allow a direct comparison of results with previous studies employing the same methodology investigating the effects of traumatic events (Back et al., 2010; Cohn et al., 2004; Keegan, 2011). Second, the hierarchical organization of the dictionary categories related to psychological processes allowed to study distinct consequences of emotional upheavals on affective, cognitive and social processes previously investigated in theoretical and empirical research (for example, Baum, 1987; Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Mehl & Pennebaker, 2003; Pennebaker et al., 2003; Pyszczynski et al., 2003; Rimé et al., 1998; Stroebe et al., 2001). Third, word count approaches such as LIWC, despite being limited and less powerful than other techniques allowing for more sophisticated reasoning on language (as discussed in Chapter 4), provide reasonable and easy-to-use tools for psychological research (Mehl, 2006).

In the analysis of Wikipedia pages, consistently with past findings on the psychological outcomes of emotional upheavals, LIWC detected a higher presence of words expressing negative emotions, cognitive activity and social processes in articles about traumatic events. This result confirmed LIWC as a valuable tool for effectively detecting expected characteristics of the collective representation of traumatic events. Specifically, the increased use of negatively connoted affective language reflected the emotional charge of traumatic events, while the use of insightful, causal and tentative words, expressing cognitive activity, suggested an active sensemaking process in the attempt to find plausible explanations. On the other hand, the higher presence of words referring to human beings and to family confirmed the increased social orientation found in the aftermath of traumatic events, while the increased presence of language expressing positive emotions and referring to friends reflected the informal context of articles about non traumatic events. The different temporal focus found in articles about old and recent traumatic events, showing an increased

presence of past tense in articles about old emotional upheaval and an increased use of present and future tenses in those about recent traumatic events confirmed previous results by Keegan (2011) on breaking news articles. Moreover, it confirmed Assmann's (1995) theoretical distinction between communicative and cultural memory, according to which collective memory is formed through the gradual passage from everyday interactions located in the present, to a more stable and objectified representation of the past. These results supported LIWC as a valuable tool for detecting different representations of articles about different types of events, allowing to further investigate collective memory formation processes in Wikipedia.

What are the psychological processes embedded in Wikipedia collective representations of natural and human-made disasters?

Following the theoretical characterization proposed by Baum (1987), according to which accidents caused by human agencies may provoke more insidious psychological effects than disasters with natural causes, Chapter 5 analyzed the language used in Wikipedia articles about natural and human-made traumatic events. The results showed that articles about human-made accidents contained significantly higher percentages of emotional language expressing anxiety and anger, and of words expressing cognitive activity and social processes. On the contrary, articles about natural disasters were characterized by an increased presence of words referring to sadness. Consistently with Baum's (1987) characterization, the unexpected loss of control typical of man-made disasters may trigger more anxiety and focused anger, because they are perceived as potentially preventable, while there is typically no expectation of control over natural events. For this reason, human and natural accidents may engender negative feelings with opposite levels of arousal, such as anger and anxiety for the former and sadness for the latter. For a similar reason, the need to find a convincing explanation may be more pressing for human-made traumatic events, whose articles reflected higher critical thinking processes and cognitive complexity in their language. Moreover, the increased presence of words indicating psychological inhibition in articles about man-made emotional upheavals supported theoretical suggestions about more severe, and insidious psychological effects of these events. Language expressing social processes was also found to be significantly higher in articles about human-made disasters. While this result, on the one hand, may confirm past research findings about an increased social orientation after emotional upheavals, which may be amplified in the case of more traumatizing experiences, on the other hand it could simply hint at a higher number of references to the human agencies at the origin of these disasters.

How do different psychological processes evolve over time during the first days following traumatic events in collaborative environments such as Wikipedia talk pages?

The last research question was addressed in Chapter 6. The analysis of the language used in the talk pages about the "7 July 2005 London bombings" and the "2011-2012 Egyptian revolution" partially confirmed previous research suggesting an increase of affective, cognitive and social processes in the aftermath of emotional upheavals (Cohn et al., 2004; Davis & Nolen-Hoeksema, 2001; Koss & Kilpatrick, 2001; Pyszczynski et al., 2003; Rimé et al., 1998; Rimé et al., 1992; Stroebe et al., 2001). Specifically, affective language related to negative emotions was significantly higher in the immediate aftermath of the London bombings in the associated talk page, and during the days of the "Battle of the Camel" in the talk page about the

Egyptian revolution. The related comments expressed grief and mourning for the victims, but also anxiety and anger in the case of the London bombings, and reflected the tense situation through the heated discussions among users in the case of the Egyptian revolution. One unexpected result with regard to previous findings by Cohn and colleagues (2004) but consistent across the two studies in Wikipedia was related to the specular temporal evolution of language expressing affective and cognitive processes. While previous research suggested a parallel temporal evolution of these psychological processes, the analysis of Wikipedia talk pages showed that in both cases when emotional language was higher, words expressing cognitive processes were lower, and when emotional language decreased, the use of insightful and discrepancy words increased. One possible explanation advanced in Chapter 6 is that, while Cohn and colleagues (2004) analyzed personal journal entries written in individual contexts, Wikipedia talk pages are especially designed to foster users' interactions and coordination activities. In this collaborative environment, the social context may influence the temporal evolution of emotional sharing and other psychological processes, intensifying conflicting arguments and inhibiting at the same time critical thinking and linguistic politeness. Finally, language related to social processes was found to be significantly higher during the first days after the London bombings in the associated talk page, and during the days of the "Battle of the Camel" in the talk page about the Egyptian revolution. In the case of the London bombings the increased social orientation was mainly related to users' concerns for helping London citizens finding information about friends and families, while in the case of the Egyptian revolution talk page, it referred to an exacerbation of the debates among editors and to the preoccupation for the journalists being harassed and for the role of women in the uprisings.

In conclusion, collective memory building is a long, dynamic and complex process which often takes several years and generations and involves many actors, both individual and collective ones, and therefore a thorough investigation of how collective memories are formed is an ambitious task. The studies presented in this dissertation constituted a first step toward the study of collective memory processes in the Web 2.0 era. The contemporary participatory culture allows people to actively produce content, and the massive backup into digital archives allows researchers to unobtrusively access textual data about people's thought and feelings, providing new research opportunities toward quantitative and longitudinal words on a large scale, which could have been difficult to carry out just few years ago. In this context, the collaborative environment of Wikipedia is particularly suitable to investigate collective, social and psychological processes involved in memory formation. In the long term, understanding the formation of collective memories through the public discourse could provide insights into the functioning of the current society. In the short term, studying individuals' reactions during the immediate aftermath of traumatizing events could allow a better comprehension of the extent of trauma and of the psychological outcomes of different emotional upheavals. Considering the importance of collective memories research for understanding our society, I believe the results presented in this dissertation can pave the way to additional investigations of this relevant topic.

Bibliography

Abel, F., Gao, Q., Houben, G.-J., & Tao, K. (2011). *Analyzing temporal dynamics in Twitter profiles for personalized recommendations in the social Web*. Paper presented at the Proceedings of ACM WebSci '11, 3rd International Conference on Web Science, Koblenz, Germany.

Adler, A. (1943). Neuropsychiatric complications in victims of Boston's Cocoanut Grove disaster. *JAMA: The Journal of the American Medical Association, 123*(17), 1098-1101. doi: 10.1001/jama.1943.02840520014004

Agerri, R., & Garcia-Serrano, A. (2010). *Q-WordNet: Extracting polarity from WordNet senses*. Paper presented at the 7th International Conference on Language Resources and Evaluation (LREC 2010), Malta.

Agirre, E., & Edmonds, P. (2006). *Word sense disambiguation*. Dordrecht, The Netherlands: Springer.

Agirre, E., & Martinez, D. (2000). *Exploring automatic word sense disambiguation with decision lists and the web*. Paper presented at the COLING Workshop on Semantic Annotation and Intelligent Content, Saarbrücken, Germany

Agirre, E., & Rigau, G. (1996). *Word sense disambiguation using Conceptual Density*. Paper presented at the Proceedings of the 16th conference on Computational linguistics - Volume 1, Copenhagen, Denmark.

Alexander, J. C., Eyerman, R., Giesen, B., Smelser, N. J., & Sztompka, P. (Eds.). (2004). *Cultural trauma and collective identity*. Berkeley, California: University of California Press.

AlJazeera. (2011). Timeline: Egypt's revolution. A chronicle of the revolution that ended the three-decade-long presidency of Hosni Mubarak. Retrieved 1 October, 2012, from http://www.aljazeera.com/news/middleeast/2011/01/201112515334871490.html

Arthur, P. L. (2008). Pixelated memory: New media responses to trauma and crisis. *IM Interactive Media, 4*, 1-19.

Assmann, J. (1995). Collective memory and cultural identity. *New German Critique, 65*, 125-133.

Baccianella, A. E. S., & Sebastiani, F. (2010, May). *SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining*. Paper presented at the Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta.

Back, M. D., Küfner, A. C. P., & Egloff, B. (2010). The emotional timeline of september 11, 2001. *Psychol Sci.* doi: 10.1177/0956797610382124

Back, M. D., Küfner, A. C. P., & Egloff, B. (2011). "Automatic or the people?". *Psychol Sci, 22*(6), 837-838. doi: 10.1177/0956797611409592

Bantum, E. O., & Owen, J. E. (2009). Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological Assessment, 21*(1), 79-88. doi: 10.1037/a0014643

Barnier, A. J., & Sutton, J. (2008). From individual to collective memory: Theoretical and empirical perspectives. *Memory, 16*(3), 177-182. doi: 10.1080/09541440701828274

Barnier, A. J., Sutton, J., Harris, C. B., & Wilson, R. A. (2008). A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research, 9*(1-2), 33-51.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* London: Cambridge University Press.

Baum, A. (1987). Toxins, technology, and natural disasters. In G. R. Van den Bos & B. K. Bryant (Eds.), *Cataclysms, crises, and catastrophes* (pp. 5-54). Washington, DC: American Psychological Association.

Baum, A., Fleming, R., & Davidson, L. M. (1983). Natural disaster and technological catastrophe. *Environment and Behavior, 15*(3), 333-354. doi: 10.1177/0013916583153004

Baum, A., Fleming, R., & Reddy, D. M. (1986). Unemployment stress: Loss of control, reactance and learned helplessness. *Social Science & Medicine, 22*(5), 509-516. doi: 10.1016/0277-9536(86)90016-x

Baum, A., & Gatchel, R. J. (1981). Cognitive determinants of reaction to uncontrollable events: Development of reactance and learned helplessness. *J Pers Soc Psychol, 40*(6), 1078-1089. doi: 10.1037/0022-3514.40.6.1078

BBC. (2011). Egypt's revolution: Interactive map Retrieved 1 October, 2012, from http://www.bbc.co.uk/news/world-middle-east-12327995

Beaumont, P. (2011, Friday 25 February 2011). The truth about Twitter, Facebook and the uprisings in the Arab world, *The Guardian.*

Blight, G., Pulham, S., & Torpey, P. (2012, 5 January 2012). Arab spring: an interactive timeline of Middle East protests, *The Guardian.* Retrieved from http://www.guardian.co.uk/world/interactive/2011/mar/22/middle-east-protest-interactive-timeline

Bohannon, J. N. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition, 29*(2), 179-196.

Booth, R. J., Petrie, K. J., & Pennebaker, J. W. (1997). Changes in circulating lymphocyte numbers following emotional disclosure: evidence of buffering? *Stress Medicine, 13*(1), 23-29. doi: 10.1002/(sici)1099-1700(199701)13:1<23::aid-smi714>3.0.co;2-e

Bowker, G. C. (2005). *Memory practices in the sciences.* Cambridge: MIT Press.

boyd, d. m., & Ellison, N. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1). doi: citeulike-article-id:1906496

Britannica Online Encyclopedia (2011). From http://www.britannica.com/

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*(1), 73-99.

Brown, S., & Hoskins, A. (2010). Terrorism in the new memory ecology: mediating and remembering the 2005 London Bombings. *Behavioural Sciences of Terrorism and Political Aggression, 2*(2), 87-107.

Bryant, S., Forte, A., & Bruckman, A. (2005). *Becoming Wikipedian: Transformation of participation in a collaborative online enciclopedia.* Paper presented at the 2005 international ACM SIGGROUP conference on Supporting group work (GROUP '05).

Bucci, W. (1995). The power of the narrative: a multiple code account. In J. W. Pennebaker (Ed.), *Emotion, Disclosure, and Health* (pp. 93-122). Washington, DC: Am. Psychol. Assoc.

Burns, A., & Eltham, B. (2009). *Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 election crisis.* Paper presented at the Communications Policy & Research Forum 2009, Sidney.

Castells, M. (2001). *The Internet galaxy: Reflections on the Internet, business, and society.* New York: Oxford University Press.

Caulfield, P. (2011, 1 February 2011). Google executive, Wael Ghonim, vanishes in Cairo. Family desperate after Twitter account goes silent, *New York Daily News.*

Chebib, N. K., & Sohail, R. M. (2011). The reasons social media contributed to the 2011 Egyptian revolution. *International Journal of Business Research and Management, 2*(39), 139-162.

Christensen, A. J., Edwards, D. L., Wiebe, J. S., Benotsch, E. G., McKelvey, L., Andrews, M., & Lubaroff, D. M. (1996). Effect of verbal self-disclosure on natural killer cell activity: moderating influence of cynical hostility. *Psychosom Med, 58*(2), 150-155.

Chung, C., & Pennebaker, J. W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social Communication.* (pp. 343-359). New York, NY US: Psychology Press.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. [Research Support, U.S. Gov't, P.H.S.]. *Psychol Sci, 15*(10), 687-693. doi: 10.1111/j.0956-7976.2004.00741.x

Conway, M. (1995). *Flashbulb memories.* Hove: Lawrence Erlbaum.

Cuc, A., Ozuru, Y., Manier, D., & Hirst, W. (2006). On the formation of collective memories: The role of a dominant narrator. *Mem Cognit, 34*(4), 752-762.

Davis, C. G., & Nolen-Hoeksema, S. (2001). Loss and meaning. *American Behavioral Scientist, 44*(5), 726-741. doi: 10.1177/0002764201044005003

Dorr, B. J., & Katsova, M. (1998). *Lexical selection for cross-language applications: combining LCS with WordNet.* Paper presented at the Proceedings of the Third Conference of the Association for Machine Translation in the Americas on Machine Translation and the Information Soup.

Eisenberger, N. I., Kemeny, M. E., & Wyatt, G. E. (2003). Psychological inhibition and CD4 T-cell levels in HIV-seropositive women. *J Psychosom Res, 54*(3), 213-224. doi: 10.1016/s0022-3999(02)00473-7

Ekeocha, J. O., & Brennan, S. E. (2008). Collaborative recall in face-to-face and electronic groups. [Research Support, Non-U.S. Gov't]. *Memory, 16*(3), 245-261. doi: 10.1080/09658210701807480

Esuli, A., & Sebastiani, F. (2006). *SENTIWORDNET: A publicly available lexical resource for opinion mining.* Paper presented at the 5th Conference on Language Resources and Evaluation.

Eyerman, R. (2004). Cultural trauma: Slavery and the formation of African American identity. In J. C. Alexander, R. Eyerman, B. Giesen, N. J. Smelser & P. Sztompka (Eds.), *Cultural Trauma and Collective Identity* (pp. 61-111). Berkeley: University of California Press.

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database.* Cambridge, MA: MIT Press.

Ferron, M. (in press). Wikimemories: La costruzione della memoria collettiva del terremoto dell'Aquila in Wikipedia. In A. Micalizzi & M. Farinosi (Eds.), *Netquake. Media digitali e disastri naturali.* Milano: Franco Angeli.

Ferron, M., & Massa, P. (2011a). *Collective memory building in Wikipedia: the case of North African uprisings.* Paper presented at the Proceedings of the 7th International Symposium on Wikis and Open Collaboration, Mountain View, California.

Ferron, M., & Massa, P. (2011b). *Studying collective memories in Wikipedia.* Paper presented at the 3rd Digital Memories Conference, Prague.

Finkenauer, C., Gisle, L., & Luminet, O. (1997). When individual memories are socially shaped: Flashbulb memories of sociopolitical events. In J. W. Pennebaker, D. Paez & B. Rimé (Eds.), *Collective Memory of Political Events: Social Psychological Perspectives* (pp. 191-208). Mahwah: Lawrence Erlbaum Associates.

Ford, H. (2012). Wikipedia sources: Managing sources in rapidly evolving global news articles on the English Wikipedia.

Forte, A., & Bruckman, A. (2005). *Why do people write for Wikipedia? Incentives to contribute to open-content publishing.* Paper presented at the GROUP 05 workshop: Sustaining community: The role and design of incentive mechanisms in online systems, Sanibel Island, FL.

Frijda, N. H. (1997). Commemorating. In J. W. Pennebaker, D. Paez & B. Rimé (Eds.), *Collective Memory of Political Events: Social Psychological Perspectives* (pp. 103-127). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Garde-Hanse, J., Hoskins, A., & Reading, A. (2009). *Save as...Digital memories.* Basingstoke: Palgrave Macmillan.

Giles, G. (2005). Internet encyclopaedias go head to head. *Nature, 438,* 900-901.

Gleser, G. C., Green, B. L., & Winget, C. (1981). *Prolonged psychosocial effects of disaster: A study of Buffalo Creek.* New York: Academic Press.

Gleser, G. C., Green, B. L., & Winget, C. N. (1978). Quantifying interview data on psychic impairment of disaster survivors. *Journal of Nervous and Mental Disease, 166*(3), 209-216. doi: 10.1097/00005053-197803000-00007

Goodwin, R., & Gaines, S. (2009). Terrorism perception and its consequences following the 7 July 2005 London bombings. *Behavioral Sciences of Terrorism and Political Aggression, 1*(1), 50-65.

Gottschalk, L. A. (1995). *Content analysis of verbal behavior: New findings and clinical applications.* Hillsdale, NJ: Erlbaum.

Gottschalk, L. A., & Bechtel, R. (1993). *Computerized content analysis of natural language or verbal texts.* Palo Alto, CA: Mind Garden.

Gottschalk, L. A., & Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior.* Berkeley: University of California Press.

Gottschalk, L. A., Gleser, G. C., Daniels, R., & Block, S. (1958). The speech patterns of schizophrenic patients: A method of assessing relative degree of personal disorganization and social alienation. *Journal of Nervous and Mental Disease, 127,* 153-166.

Gottschalk, L. A., Winget, C. N., & Gleser, G. C. (1969). *Manual of instructions for using the Gottschalk-Gleser content analysis scales: Anxiety, hostility, social alienation-personal disorganization*. Berkeley: University of California Press.

Green, B. L. (1980). *Prediction of long-term psycho-social functioning following the Beverly Hills fire*. University of Cincinnati.

Green, B. L. (1996). Traumatic stress and disaster: Mental health effects and factors influencing adaptation. In F. L. Mak & C. C. Nadelson (Eds.), *International review of psychiatry, Vol. 2*. (pp. 177-210). Washington, DC US: American Psychiatric Association.

Grossman, L. (2009, 17 June 2009). Iran's protests: Why Twitter is the medium of the movement, *Time*.

Halbwachs, M. (1950). *The collective memory*. New York: Harper Colophon Books.

Halbwachs, M. (1992). *On collective memory*. Chicago: University of Chicago Press.

Hara, N., Shachaf, P., & Hew, K. F. (2010). Cross-cultural analysis of the Wikipedia community. *Journal of the American Society for Information Science and Technology, 61*(10), 2097-2108. doi: 10.1002/asi.21373

Harris, C. B., Paterson, H. M., & Kemp, R. I. (2008). Collaborative recall and collective memory: What happens when we remember together? *Memory, 16*(3), 213-230. doi: 10.1080/09658210701811862

Hart, R. P. (1984). *Verbal style and the presidency: A computer based analysis*. New York: Academic Press.

Hart, R. P. (2001). Redeveloping DICTION: Theoretical considerations. In M. D. West (Ed.), *Theory, method, and practice in computer content analysis* (pp. 43-60). New York: Ablex.

Henderson, S., & Bostock, T. (1977). Coping behaviour after shipwreck. *The British Journal of Psychiatry, 131*, 15-20. doi: 10.1192/bjp.131.1.15

Hilton, A., & Armstrong, R. A. (2006). Statnote 6: post-hoc ANOVA tests. *Microbiologist*, 34-36.

Hirst, W., & Manier, D. (2008). Towards a psychology of collective memory. [Review]. *Memory, 16*(3), 183-200. doi: 10.1080/09658210701811912

History Central (2011). From http://www.historycentral.com/

Hofstede, G. H. (1991). *Cultures and organizations: Software of the mind*. London: McGraw-Hill.

Horowitz, M. (1976). *Stress response syndromes*. Newy York: Aronson.

Hoskins, A. (2009a). Digital network memory. In A. Erll & A. Rigney (Eds.), *Mediation, Remediation, and the Dynamics of Cultural Memory* (pp. 91-108). Berlin: Mouton de Gruyter.

Hoskins, A. (2009b). The mediatization of memory. In J. Garde-Hanse, A. Hoskins & A. Reading (Eds.), *Save As...Digital Memories*. Basingstoke: Palgrave Macmillan.

Information Britain (2011). From http://www.information-britain.co.uk/

Irwin-Zarecka, I. (1994). *Frames of remembrance*. New Brunswick: Transation Publishers.

Jedlowski, P. (2001). Memoria, mutamento sociale e modernità *La memoria contesa. Studi di comunicazione sociale del passato* (pp. 40-67). Milan, Italy: Franco Angeli.

Jurkowitz, M. (2011). Events in Egypt trigger record coverage: PEJ News Coverage Index.

Keegan, B. (2011). *Breaking news, breaking planes, and breaking hearts: Psycholinguistics and sensemaking in collaborative accounts of catastrophe.* Paper presented at the International Communication Association, Boston, MA.

Keegan, B., Gergle, D., & Contractor, N. (2011). *Hot off the wiki: Dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes.* Paper presented at the Proceedings of the 7th International Symposium on Wikis and Open Collaboration, Mountain View, California.

Khamis, S., & Vaughn, K. (2011). Cyberactivism in the Egyptian revolution: How civic engagement and citizen journalism tilted the balance. *Arab Media and Society, 13.*

Kittur, A., Suh, B., Pendleton, B. A., & Chi, E. H. (2007). *He says, she says: conflict and coordination in Wikipedia.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.

Knight, K. (1993). *Building a large ontology for machine translation.* Paper presented at the Proceedings of the workshop on Human Language Technology, Princeton, New Jersey.

Koss, M. P., & Kilpatrick, D. G. (2001). Rape and sexual assault. In E. Gerrity, T. M. Keane & F. Tuma (Eds.), *The mental health consequences of torture* (pp. 177-193). New York: Kluwer Academic/Plenum.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261-273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Langer, S., & Hickey, M. (1998). Using semantic lexicons for full text message retrieval in a communication aid. *Nat. Lang. Eng., 4*(1), 41-55. doi: 10.1017/s1351324998001855

Laniado, D., Tasso, R., Volkovich, Y., & Kaltenbrunner, A. (2011). *When the Wikipedians talk: Network and tree structure of Wikipedia discussion pages.* Paper presented at the ICWSM 2011 - 5th International AAAI Conference on Weblogs and Social Media.

Leopold, L. R., & Dillon, H. (1963). Psycho-anatonomy of a disaster: A long term study of post-traumatic neurosis in survivors of a marine explosion. *American Journal of Psychiatry, 119*, 913-921.

Lutgendorf, S. K., & Antoni, M. H. (1999). Emotional and cognitive processing in a trauma disclosure paradigm. *Cognitive Therapy and Research, 23*(4), 423-440. doi: 10.1023/a:1018760118863

Magnini, B., & Cavaglià, G. (2000). *Integrating subject field codes into wordnet.* Paper presented at the LREC-2000, the Second International Conference on Language Resources and Evaluation, Athens, Greece.

Magnini, B., Negri, M., Prevete, R., & Tanev, H. (2002). *A WordNet-based approach to Named Entities recognition.* Paper presented at the Proceedings of the 2002 workshop on Building and using semantic networks - Volume 11.

Magnini, B., & Strapparava, C. (2001). *Using WordNet to improve user modelling in a web document recommender system.* Paper presented at the NAACL 2001 Workshop on WordNet and Other Lexical Resources, Pittsburgh, PA.

Maher, H., & Eskandar, W. (2012). Timeline: Egypt's year of revolution, *Ahram Online.* Retrieved from
http://english.ahram.org.eg/NewsContent/1/114/32352/Egypt/-January-Revolution-continues/Timeline-Egypts-year-of-revolution.aspx

Mandala, R., Tokunaga, T., & Tanaka, H. (1998). The Use of WordNet in information retrieval. In S. Harabagiu (Ed.), *Use of Word{N}et in Natural Language Processing Systems: Proceedings of the Conference* (pp. 31-37): Association for Computational Linguistics.

Maniu, S., Cautis, B., & Abdessalem, T. (2011). *Building a signed network from interactions in Wikipedia.* Paper presented at the Databases and Social Networks, Athens, Greece.

Massa, P. (2011). *Social networks of Wikipedia.* Paper presented at the ACM Hypertext 2011 Conference.

McGuire, L., & Kiecolt-Glaser, J. K. (2000). *Interpersonal pathways to health* (Vol. 63).

Mehl, M. R. (2006). Quantitative text analysis. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 141-156). Washington, DC: American Psychological Association.

Mehl, M. R., & Pennebaker, J. W. (2003). The social dynamics of a cultural upheaval: social interactions surrounding September 11, 2001. [Research Support, U.S. Gov't, P.H.S.]. *Psychol Sci, 14*(6), 579-585.

Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology, 64*, 1306-1315.

Micalizzi, A. (2012). Traumatic event and digital memories: Remembering and processing the earthquake in Abruzzi P. Wilson & P. McEntaggart (Eds.), *Navigating Landscapes of Mediated Memory*

Middleton, D., & Brown, S. D. (2005). *The social psychology of experience: Studies in remembering and forgetting.* London: Sage.

Morozov, E. (2009). Iran: Downside to the "Twitter Revolution". *Dissent, 56*(4), 10-14.

Motulsky, H. (1999). *Analyzing data with GraphPad Prism.* San Diego, CA: GraphPad Software, Inc.

Neal, A. (1998). *National trauma and collective memory: Major events in the American century.* Armonk: M.E. Sharpe.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying Words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665-675. doi: 10.1177/0146167203029005010

Nora, P. (1989). Between memory and history: Les lieux de mémoire. *Representations 26*, 7-24.

Olick, J. K. (2008). 'Collective memory': A memoir and prospect. *Memory Studies, 1*(1), 23-29.

Olick, J. K., & Levy, D. (1997). Collective memory and cultural constraint: Holocaust myth and rationality in German politics. *American Sociological Review, 62*(6), 921-936.

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation, 15*(12), 1-9.

Owais, R. (2011). Arab media during the Arab spring in Egypt and Tunisia: Time for change. *Middle East Media Educator, 1*(1), 9-13.

Paez, D., Basabe, N., & Gonzalez, J. L. (1997). Social processes and collective memory: A cross cultural approach to remembering. In J. W. Pennebaker, D. Paez & B. Rimé (Eds.), *Collective Memory of Political Events: Social Psychological*

Perspectives (pp. 147-174). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Pennebaker, J. W. (1993). Putting stress into words: Health, linguistic, and therapeutic implications. [Research Support, U.S. Gov't, Non-P.H.S.]. *Behav Res Ther, 31*(6), 539-548.

Pennebaker, J. W. (1997a). *Opening up: The healing power of emotional expression*. New York: Guilford.

Pennebaker, J. W. (1997b). Writing about emotional experiences as a therapeutic process. *Psychol Sci, 8*(3), 162-166. doi: 10.1111/j.1467-9280.1997.tb00403.x

Pennebaker, J. W., & Banasik, B. L. (1997). On the creation and mantainance of collective memories: History as social psychology. In J. W. Pennebaker, D. Paez & B. Rimé (Eds.), *Collective Memory of Political Events: Social Psychological Perspectives* (pp. 3-19). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Pennebaker, J. W., Barger, S. D., & Tiebout, J. (1989). Disclosure of traumas and health among Holocaust survivors. [Research Support, U.S. Gov't, Non-P.H.S.]. *Psychosom Med, 51*(5), 577-589.

Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. [Research Support, U.S. Gov't, P.H.S.]. *J Abnorm Psychol, 95*(3), 274-281.

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10*(6), 601-626. doi: 10.1080/026999396380079

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah: Lawrence Erlbaum Associates.

Pennebaker, J. W., & Harber, K. D. (1993). A social stage model of collective coping: The Loma Prieta earthquake and the Persian Gulf War. *Journal of Social Issues, 49*(4), 125-145. doi: 10.1111/j.1540-4560.1993.tb01184.x

Pennebaker, J. W., & Lay, T. C. (2002). Language use and personality during crises: Analyses of mayor Rudolph Giuliani's press conferences. *Journal of Research in Personality, 36*(3), 271-282.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *J Pers Soc Psychol, 72*(4), 863-871.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annu Rev Psychol, 54*, 547-577. doi: 10.1146/annurev.psych.54.101601.145041

Pennebaker, J. W., Paez, D., & Rimé, B. (1997). *Collective memory of political events: Social psychological perspectives*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Pentzold, C. (2009). Fixing the floating gap: The online encyclopaedia Wikipedia as a global memory place. *Memory Studies, 2*(2), 255-272. doi: 10.1177/1750698008102055

Petrie, K. J., Booth, R. J., & Pennebaker, J. W. (1998). The immunological effects of thought suppression. *J Pers Soc Psychol, 75*(5), 1264-1272. doi: citeulike-article-id:277963

Pfeil, U., Zaphiris, P., & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication, 12*(1), 88-113. doi: 10.1111/j.1083-6101.2006.00316.x

Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt on President Reagan. *Cognition, 16*(1), 63-80.

Ploeger, A. (1977). A 10-year follow up of miners trapped for 2 weeks under threatening circumstances. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety: IV*. (pp. 23-28). Oxford England: Hemisphere.

Pury, C. L. S. (2011). Automation can lead to confounds in text analysis: Back, Küfner, and Egloff (2010) and the not-so-angry Americans. *Psychol Sci, 22*(6), 835-836. doi: 10.1177/0956797611408735

Pyszczynski, T., Solomon, S., & Greenberg, J. (2003). *In the wake of 9/11: The psychology of terror*. Washington, DC US: American Psychological Association.

Recumber, T. (2012). The prosumption of commemoration. *American Behavioral Scientist, 56*(4), 531-549. doi: 10.1177/0002764211429364

Reser, J. P. (2007). The Experience of Natural Disasters: Psychological Perspectives and Understandings. . In J. Lidstone, L. M. Dechano & J. P. Stoltman (Eds.), *International Perspectives on Natural Disasters: Occurrence, Mitigation, and Consequences* (Vol. 21, pp. 369-384): Springer Netherlands.

Rimé, B. (1995). Mental rumination, social sharing and the recovery from emotional exposure. In J. W. Pennebaker (Ed.), *Emotion, disclosure and health* (pp. 271-292). Washington, DC: American Psychological Association.

Rimé, B., & Christophe, V. (1997). How individual emotional episodes feed collective memory. In J. W. Pennebaker, D. Paez & B. Rimé (Eds.), *Collective Memory of Political Events: Social Psychological Perspectives* (pp. 131-146). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Rimé, B., Finkenauer, C., Luminet, O., Zech, E., & Philippot, P. (1998). Social sharing of emotion: New evidence and new questions. *European Review of Social Psychology, 9*(1), 145-189. doi: 10.1080/14792779843000072

Rimé, B., Philippot, P., Boca, S., & Mesquita, B. (1992). Long lasting and social consequences of emotions: Social sharing and rumination. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 3, pp. 225-258). Chichester: Wiley.

Rosenberg, S. D., & Tucker, G. J. (1978). Verbal behavior and schizophrenia: The semantic dimension. *Archives of General Psychiatry, 36*, 1331-1337.

Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition, 16*(1), 81-95.

Russell, J. A. (1980). A circumplex model of affect. *J Pers Soc Psychol, 39*(6), 1161-1178. doi: 10.1037/h0077714

Sexton, J. B., & Helmreich, R. L. (2003). Using language in the cockpit: Some linguistic factors. In R. Dietrich & T. von Meltzer (Eds.), *Communication in High Risk Environments* (pp. 57-74). Hamburg: Helmut Buske Verlag.

Shandler, J. (1999). *While America watches: Televising the Holocaust*. New York: Oxford University Press.

Sheskin, D. J. (2004). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton, FL: Chapman & Hall.

Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., & Dunne, C. (2010). NodeXL: A free and open network overview, discovery and exploration add-in for Excel 2007/2010. Retrieved from http://nodexl.codeplex.com/

Smyth, J. M., Stone, A. A., Hurewitz, A., & Kaell, A. (1999). Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis: A randomized trial. *JAMA: The Journal of the American Medical Association, 281*(14), 1304-1309. doi: 10.1001/jama.281.14.1304

Sokal, R. R. (1981). *Biometry*. New York: W.H. Freeman.

Stone, A. A., Turkkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (2000). *The science of self-report: Implications for research and practice.* Mahwah, NJ: Erlbaum.

Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M. (1966). *The general inquirer: A computer approach to content analysis.* Cambridge: MIT Press.

Strapparava, C., & Valitutti, A. (2004). *WordNet-Affect: An affective extension of WordNet.* Paper presented at the Proceedings of LREC.

Strapparava, C., Valitutti, A., & Stock, O. (2006). *The affective weight of lexicon.* Paper presented at the Fifth International Conference on Language Resources and Evaluation, Genoa, Italy.

Stroebe, M. S., Hansson, R. O., Stroebe, W., & Schut, H. (2001). *Handbook of bereavement research: Consequences, coping, and care.* Washington, DC US: American Psychological Association.

Sturken, M. (1997). *Tangled memories: The Vietnam War, the AIDS epidemic, and the politics of remembering.* London: University of California Press.

Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005). *Assessing information quality of a community-based encyclopedia.* Paper presented at the International Conference on Information Quality - ICIQ 2005.

Sztompka, P. (2000). Cultural trauma. *European Journal of Social Theory, 3*(4), 449-466. doi: 10.1177/136843100003004004

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24-54. doi: 10.1177/0261927x09351676

Titchener, J. L., & Kapp, F. T. (1976). Family and character change at Buffalo Creek. *The American Journal of Psychiatry, 133*(3), 295-299.

Toulmin, S. E. (2003). *The uses of argument.* Cambridge: Cambridge University Press.

Van House, N., & Churchill, E. F. (2008). Technologies of memory: Key issues and critical perspectives. *Memory Studies, 1*(3), 295-310. doi: 10.1177/1750698008093795

Viégas, F., Wattenberg, M., & Dave, K. (2004). *Studying cooperation and conflict between authors with history flow visualizations.* Paper presented at the SIGCHI 2004.

Viegas, F. B., Wattenberg, M., Kriss, J., & Ham, F. v. (2007). *Talk before you type: Coordination in Wikipedia.* Paper presented at the Proceedings of the 40th Annual Hawaii International Conference on System Sciences.

von Hein, M. (2011, 11 February 2011). Timeline of the revolution in Egypt, *Deutsche Welle.* Retrieved from http://www.dw.de/dw/article/0,,14837364,00.html

Voss, J. (2005). *Measuring Wikipedia.* Paper presented at the 10th International Conference of the International Society for Scientometrics and Informetrics.

Wang, Q. (2008). On the cultural constitution of collective memory. *Memory, 16*(3), 305-317. doi: 10.1080/09658210701801467

We Are All Khaled Said (2010). In Facebook. Retrieved 1 October 2012, from http://www.facebook.com/elshaheeed.co.uk

Weik, M. (1996). *Communications standard dictionary.* New York: Chapman & Hall.

Weintraub, W. (1981). *Verbal behavior: Adaptation and psychopathology.* New York: Springer.

Weintraub, W. (1989). *Verbal behavior in everyday life.* New York: Springer.

Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. [Research Support, Non-U.S. Gov't]. *J Exp Psychol Learn Mem Cogn, 23*(5), 1160-1175.

Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., & Smith, M. (2011). *Finding social roles in Wikipedia.* Paper presented at the Proceedings of the 2011 iConference, Seattle, Washington.

Welser, H. T., Underwood, P., Cosley, D., Hansen, D., & Black, L. (2010). Wiki-networks: Connections of creativity and collaboration. In D. Hansen, B. Shneiderman & M. Smith (Eds.), *Analyzing Social Media Networks with NodeXL* (pp. 247-271). New York: Morgan Kaufmann.

Wertsch, J. V. (2002). *Voices of collective remembering.* Cambridge: Cambridge University Press.

Whitlock, C., & Fadel, L. (2011). Protesters surge into Tahrir Square as Google executive energizes movement, *The Washington Post.* Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2011/02/07/AR2011020706013.html

Wikipedia: What Wikipedia is not (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not#FORUM

Wikipedia: Writing better articles (2004). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:1SP

Wikipedia:Bots (2007). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Bots

Wikipedia:Etiquette (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Etiquette

Wikipedia:Featured articles (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Featured_articles

Wikipedia:Neutral point of view (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view

Wikipedia:Size of Wikipedia (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia

Wikipedia:Staying cool when the editing gets hot (2002). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Staying_cool_when_the_editing_gets_hot

Wikipedia:Talk page guidelines (2005). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:Talk_page_guidelines

Wikipedia:What Wikipedia is not (2001). In Wikipedia. Retrieved 1 October 2012, from http://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not

Wilcox, R. R. (2009). *Basic statistics. Understanding conventional methods and modern insights.* New York: Oxford University Press.

Winograd, E., & Neisser, U. (1992). *Affect and accuracy in recall: Studies of 'flashbulb' memories.* Cambridge: Cambridge University Press.

Wolman, D. (2012, 10 February 2012). The digital road to Egypt's revolution, *The New York Times.* Retrieved from http://www.nytimes.com/interactive/2012/02/12/opinion/sunday/20120212-tahir-timeline.html

Wortman, C. B., & Brehm, J. W. (1975). Responses to Uncontrollable Outcomes: An Integration of Reactance Theory and the Learned Helplessness Model. In

B. Leonard (Ed.), *Advances in Experimental Social Psychology* (Vol. Volume 8, pp. 277-336): Academic Press.

Zerubavel, E. (2003). *Time maps: Collective memory and the social shape of the past.* Chicago: University of Chicago press.

Zerubavel, Y. (1995). *Recovered roots: Collective memory and the making of Israeli national tradition.* Chicago: University of Chicago press.

Zhu, H., Kraut, R. E., Wang, Y.-C., & Kittur, A. (2011). *Identifying shared leadership in Wikipedia.* Paper presented at the PART 5 -------- Proceedings of the 2011 annual conference on Human factors in computing systems, Vancouver, BC, Canada.

Zickuhr, K., & Rainie, L. (2011). Wikipedia, past and present: Pew Internet and American Life Project.