Online Together

A Sociological Study of the Concept of Togetherness and the Contemporary Conditions for Social Interaction

Lovisa Eriksson
Abstract


The recent advances in digital communication technologies have altered the way in which people socialize on a day-to-day basis. A question that has arisen in relation to this is what being somewhere together actually means at a time when our interactions are no longer confined to shared physical places.

The phenomenon of being somewhere together (also: togetherness) has previously been studied within the fields of social presence theory (which focuses on digitally mediated ‘togetherness’ and primarily departs from a psychological perspective) and microsociology (which takes an arguably more interactional approach to the idea of being together but primarily focuses on face-to-face interaction). Therefore, what is missing is a conceptualization of togetherness that can account both for togetherness in contexts other than those mediated face-to-face and for the ways in which togetherness is potentially ‘created’ in social interaction.

The purpose of this thesis is to address this shortcoming by examining the underlying problem of being together and the conceptualizations of being together in the two aforementioned discourses. For the theoretical analyses, the example of online chat conversation is used as the primary focus of study.

The thesis comprises three main parts. In the first part, the question of why being together has become difficult to conceptualize since the introduction of electronic and digital communication technologies is explored. The second part of the thesis is a review of what being together stands for in social presence theory and microsociology, respectively. In the third part, the two reviewed understandings of being together are examined. Here, it is observed that social presence theory portrays being together as something that occurs in informational environments, while microsociology portrays it instead as something pertaining to framed (or specified) social situations. Thereafter follows a critical examination of being together in informational situations and being together in framed social situations in which the notions are analysed in relation to online chat. It is concluded that the second view of being together (as a framed activity) is more promising for the future study of togetherness in online chat environments, and potentially also for togetherness in digitally mediated environments more generally.

Keywords: Togetherness, Social Presence, Computer-Mediated Communication, Microsociology

Lovisa Eriksson, Department of Sociology, Box 624, Uppsala University, SE-75126 Uppsala, Sweden.

© Lovisa Eriksson 2016

ISSN 0585-5551
ISBN 978-91-554-9593-0
urn:nbn:se:uu:diva-285860 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-285860)
# Table of Contents

1. Introduction ................................................................................................................................. 9  
   Being Together in Two Fields of Research .................................................................................. 12  
   Social Presence Research ........................................................................................................... 13  
   The Sociological Study of Social Meetings ............................................................................... 15  
   The Example of Online Chat Conversations ........................................................................... 17  
   Purpose and Research Questions ............................................................................................... 21  
   Methodology and Outline ........................................................................................................... 23  

2. Being Where Together? ................................................................................................................. 26  
   Environments, Situations, and Social Situations ........................................................................ 27  
   Physical Environments and Situations ....................................................................................... 28  
   The Different Layers of the Social Situation ............................................................................... 30  
   Environments, Situations, and New Communication Technologies ... 32  
   The Non-Spatiality of the Digitally Mediated Environment ................................................ 35  
   The Digitally Mediated Environment as an Embedded Environment ................................ 38  
   The Medium as an Informational Environment ...................................................................... 42  
   Informational Affordances of Physical and Digitally Mediated Environments ....................... 43  
   General Affordances of ‘Pure’ Physical Environments ....................................................... 44  
   Digitally Mediated Environments and their Affordances .................................................... 47  
   Discussion and Summary .......................................................................................................... 50  

3. Being (T/Here) Together ............................................................................................................. 55  
   Social Presence Theory and ‘Being There Together’ ............................................................... 55  
   Social Presence as an Environmental Affordance ................................................................... 56  
   Social Presence as a Quality of the Medium ........................................................................... 56  
   Social Presence as Observability ............................................................................................. 57  
   Social Presence as Self-Presentation ....................................................................................... 58  
   Social Presence as an Experience of External Events ............................................................. 59  
   Social Presence as an Experience of the Medium ................................................................. 60  
   Social Presence as the Experience of High Social Realism or as Distorted Perception .......... 60  
   Social Presence as an Experience of Others .......................................................................... 62  
   Social Presence as an Experience of Reciprocity ................................................................... 63
The Influence of Relational or Interational Factors on the Experience of Social Presence .................................................................65
Summary and Discussion of Social Presence Theory .................................................................66
“Being Together” in Phenomenology and Interactionist Theory ........................................71
From Independency to Intendependency ...........................................................................73
Closeness and Distance in Social Experience ......................................................................80
Indirect Social Experience .................................................................................................82
Direct Social Experience ..................................................................................................86
Intensity/Involvement .......................................................................................................88
Distance Through Misinvolvement ................................................................................90
Intimacy ..............................................................................................................................92
Summary and Discussion of the Microsociological Perspective .........................................95
Discussion and Summary ..................................................................................................101

4. Co-Located versus Co-Present Togetherness: The Case of Being Together in Online Chat.................................................................................................................................................104
Social Presence Theory and Co-located Togetherness ......................................................107
Co-located Togetherness and Online Chat Environments ............................................109
Digitally Mediated Environments do not challenge Basic Social
‘Instincts’ ............................................................................................................................109
‘Real’ versus ‘Virtual’ Co-Location .................................................................................113
The Actor as a Passive Consumer of Stimuli ................................................................116
Why Togetherness Extends Beyond Co-location:
The Example of Chatbots .................................................................................................119
Microsociology and Togetherness as Co-Presence ..........................................................124
Togetherness in Online Chat from the Perspective of Co-Presence .... 127
Framework in the Online Chat Situation ......................................................................128
Getting Closer in Online Chat Situations .....................................................................137
Involvement .......................................................................................................................137
Mental Closeness ............................................................................................................144
Intimacy ..............................................................................................................................146
Chapter Summary .............................................................................................................151

5. Summary and Conclusions ............................................................................................156
References ..............................................................................................................................168
There are a number of people to thank for their support in the making of this thesis. To begin with, I would like to thank Michael Allvin for supervising my work and for providing me with valuable feedback and commentary. A heartfelt thanks goes out to Bo Lewin for being my supervisor, but more importantly for being a mentor and dear friend. I apologize in advance for all the objects with agency that slipped through the proofreading.

I am indebted to Hannah Bradby for carefully reading and commenting on both early and more recent drafts of this thesis and whose to-the-point feedback proved absolutely invaluable, especially in the later stages of the project. Thanks also to Árni Sverrisson, Iveta Jurkane-Hobein, and Kitty Lassianatti for the helpful feedback you gave me at my interim and final seminars. I am, moreover, grateful for the advice and support I’ve received over the years from Vessela Misheva, Andrew Blasko, Wendelin Reich and Ulf Hedestig.

I want to acknowledge all my colleagues and friends at the department of Sociology at Uppsala University. A special thanks goes to my roommates Kalle Berggren and Hedvig Gröndal, and to Magdalena Vieira for keeping me company during the weekend shifts. I’d also like to thank Ulrika Söderlind, Helena Olsson, Katriina Östensson and Margareta Mårtensson for making the administrative side of academic life run much smoother. Posting this thesis would have been much more cumbersome without the help of Anna-Liv Jonsson. Thanks for being my proxy.

I could not have finished this book without the unconditional love of my family—thank you for both supporting and disrupting my flow. Hedvig, Emma, Elin, Lina, and Anna Liv: I am incredibly grateful for your patience and for all the times you gave me some (much needed) perspective. Thanks also to Helena, to all my friends in London and to my new family in South Africa. Baie dankie. If he could read, this would have been the place where I also thanked Mårre the dog for reminding me that happiness is found in the little things.

I lastly want to thank Willem for his stubborn love and for never ceasing to believe in me. I am so very grateful for our kind of togetherness.
A couple of years ago, on a late January afternoon, I met Joan—a chatbot created by the mobile app developer company Icogno. After we had talked for a while, I decided to see whether Joan could help me get to the bottom of a question I had been thinking about lately.

“Joan”, I typed. “Where am I?”

“In your computer,” she promptly replied.

But however entertaining it may be to talk to computer programs, the conversation I had with Joan that day did not exactly make me feel like I was ‘in my computer’. During our chat, I was—if anywhere—on an uncomfortable chair at my kitchen table, painfully aware of dirty dishes and the dog gnawing on the carpet. Yet our conversation kept on unfolding on an Internet website ‘in my computer’.

Later that evening, a friend of mine briefed me on the latest news of her adventurous love life. Had I then asked Joan where I was, her answer would not have been a far cry from reality. As my friend and I communicated by sending instant messages over an online chat service, I was definitely more ‘in my computer’ (or in the conversation, which materialized on my computer screen) than on that kitchen chair of mine. Although I was physically located at the exact same spot during both conversations, only the latter conversation (with my friend) evoked something of an experience of being somewhere with someone, and that ‘somewhere’ was not exactly identical to the location of my body.

It is not coincidental that computer-mediated communication was chosen to serve as the opening example of this thesis. When the communication technologies available were limited to cave painting, smoke signalling, and letter writing, it is fairly safe to say that reflections on the difference between one’s physical and social whereabouts were kept to a minimum. The communication supported by these technologies was characterized by asynchro-
nous exchange of messages (i.e. where there is a significant time delay before a sent message reaches the receiver) and was not particularly easy or time efficient to access and master. This suggests that the experience of being together somewhere (hereafter also referred to as togetherness) was probably exclusive to instances where actors were interacting face-to-face, and the social places people were in thereby coincided with the physical places they shared with each other. But with the invention of (electronic and digital) communication technologies allowing information to travel over a long distance in a very short time, it has become increasingly clear that our physical locations do not necessarily determine where we are socially, and that social place is not identical to physical place.

If we return to the opening example, there is a second aspect to my experiences of talking to the chatbot versus my friend. In both conversations, the ‘place’ and ‘method’ of interaction remained the same, yet I only experienced that I was truly together with my friend. Something was missing from my encounter with the chatbot that could not be explained by the affordances of the communication technology in use (because the same type of technology was used in both conversations). In this example, one could of course blame the lacking sense of togetherness entirely on the non-humanness of the chatbot, but I believe that would be to simplify the issue.

In fact, we do not have to look far to find examples of all-human face-to-face meetings where actors feel they are not entirely ‘there’ with each other. For example, Phaedrus—the tragic hero of Pirsig’s *Zen and the Art of Motorcycle Maintenance* (1974/1999:87)—embodies this idea of the “absent” interaction partner, when “[e]ven in the presence of others he was completely alone. People sometimes felt this and felt rejected by it, and so did not like him, but their dislike was not important to him.” If we normally take for granted that individuals who are meeting each other face-to-face are both physically and socially ‘there’ together, this assumption may not be true in all cases. Imagine, for instance, that dreadful moment when you realize that your face-to-face interaction partner is—like Phaedrus—simply not ‘with you’ any longer, that you have been abandoned in conversation, so to speak, even though the physical body of the other is still standing right in front of you. Similarly, one could ask if I am really ‘there’ at a dinner party if I am mentally preoccupied by an event that happened earlier that day, and am thereby unable to keep up with the conversation? Certainly, my body is ‘there’ at the table with the others, but that does not stop my thoughts from being elsewhere, and so my behaviour will probably reflect that I am not fully participating in the conversation.

So even when social places coincide with shared physical places, this still does not seem to imply that a shared physical space always guarantees that the inhabitants of that space are socially ‘there’ together. Another way of putting it would be to say that the sense of being together is not entirely dependent on the means we use to interact, whether directly through our bodies
or through digital communication technologies. What this illustrates is that, besides giving rise to all sorts of ‘new’ questions for scholars to study, the addition of electronic (and later, digital) communication technology in everyday social life has also directed attention to questions that were previously considered too marginal or abstract to be given full attention. For instance, Joshua Meyrowitz (1985) might not have investigated the difference between social and physical places unless electronic media had drawn his attention to it. Nonetheless, he firmly argues that electronic media did not create this difference—it only accentuated it. The driving question of this dissertation follows a similar path in that it concerns what being together somewhere entails in a time when our social interactions are becoming more and more digitalized. Yet the question of what it means to be together is not brand new, nor has the phenomenon of being together emerged because of new technology (as we are not just now starting to be able to be together). Instead, digital communication technology highlights—and perhaps also complicates—the question of what it means to be together.

The relevancy of the topic of digitally mediated togetherness becomes clear when we look at how people, at least in the more economically developed parts of the world, choose to communicate with others. Today, socializing with people that we do not share a physical space with has without doubt become an integral part of everyday life. In Sweden, which is among the countries with the highest Internet penetration in the world (Internet Society 2014:130), 83 per cent of Swedes between 17 and 83 years of age reported that they accessed the Internet more or less every day in 2014 (Statistiska Centralbyrån 2014). Of those, a majority engaged in social activities online—83 per cent sent or received e-mails, 65 per cent used the Internet to communicate with people through online chat services, instant messaging, video- or phone calls, weblogs, social networking or discussion sites, and 18 per cent played online games with others.

Even though the Internet is still mostly accessed from home in Sweden, the increased availability and use of portable devices, such as smartphones and tablet computers, have enabled people to access the Internet just about anywhere (Findahl 2014). The earlier depiction of a lonely Internet user who goes online through his stationary computer in the confinements of his home must therefore be replaced by the image of a much more mobile user. As long as they have a device for it, and an Internet network for the device to connect to, individuals can now use their mobile devices to socialize with people who are not in their physical proximity, almost anywhere. So if, during the early days of the Internet, there was a more pronounced separation between our lives ‘online’ and our lives ‘offline’, this boundary has weakened over time. Taken together, this means that the time we spend interacting with people that do not share our physical place has increased, and that the number of physical places from which we are able to access those distant
‘social places’ has also increased. The digitally mediated realm has, in other words, become almost ubiquitous.

There is little doubt that the increased use and availability of digital alternatives for social interaction have changed our everyday ways of being with each other, and that these changes to social life are important to include in the sociological purview. But as was described above, the separation of social place from physical place has complicated the picture of what togetherness, or being together, actually stands for. From the discussion so far, one can for example infer that two individuals are not automatically socially ‘there’ together just because the geographical distance between their physical bodies is small, but also that two individuals are not automatically socially not ‘there’ together just because their physical bodies are very far apart, geographically speaking. So what is it, then, to be somewhere together?

**Being Together in Two Fields of Research**

There could be many potential answers to the question of what it means to be somewhere together, all depending on what perspective we depart from, what analytical level we are operating on, and what we hope to achieve once the answer is found. A philosopher might search for an ontological answer in Heidegger’s (1927) Being and Time and conclude that being there (Dasein) always assumes being with (Mitsein), that is—our being presupposes that we are being in a world shared with others. By contrast, a student of theology might approach the question of what it means to be together from the perspective of a believer’s relationship to his God. What does God’s presence entail, and what does it mean to be in the presence of God? Even within the sociological discipline, the question of what it means to be together is bound to lead to different associations depending on the interests of the researcher. A sociologist whose research interest concerns social injustice might for instance think of togetherness in terms of what social groups and categories an actor must belong to in order to be allowed to participate in different social gatherings, and what the underlying reasons for these practices of inclusion and exclusion are. She might look for answers by studying discriminating practices on both structural and individual levels.

As the opening section suggests, the research interests that have prompted this dissertation firstly concern something that occurs, or does not occur, in the meeting between two or more particular actors. Excluded is therefore the ‘meeting’ between a person and his God or the generalized other (except perhaps the extent to which these abstractions manifest themselves in encounters between particular others), and the enquiry has a more ontic than ontological character. The research interest also primarily concerns how our ‘being together’ relates to what means we use to communicate, and how we communicate given those means, rather than how we are included in or ex-
cluded from certain social encounters on the basis of our social status or identities. Insofar as the thesis considers the influence of social norms and structures on digitally mediated togetherness, the focus will be on how these norms and structures shape and are shaped by communication technology. All these things considered, the natural place to begin looking for answers to the question of what it means to be somewhere together is in studies of mediated communication and studies of social meetings.

Social Presence Research

Following the last century’s development of new technological tools for telecommunication and teleoperation\(^2\) there came a growing interest in presence, or the phenomenon of ‘being’ in a place. Today, the question of what it means to ‘be’ in a remote place, such as a virtual environment\(^1\), still attracts attention. Although the phenomenon goes under several names (e.g. spatial presence or virtual presence\(^4\)), it was initially coined “telepresence” by Marvin Minsky (1980). With this new concept, Minsky sought to highlight users’ experiences of teleoperations: “[t]elepresence emphasizes the importance of high-quality sensory feedback and suggests future instruments that will feel and work so much like our own hands that we won't notice any significant difference.” From this notion grew an entirely new field of research that had at its core an ambition to make sense of the experience of being, operating, and communicating in virtual environments.

Soon, scholars also started paying attention to the social aspect of ‘being’ in remote environments. This is no coincidence, as the use of various kinds of digital communication technology exposes how the location of people’s physical bodies does not necessarily overlap with their social ‘location’. The meaning of ‘being somewhere’ certainly depends on what kind of ‘being’ we are talking about, and digital interaction highlights that being physically present is not the same as being socially present.

In this field, it is assumed that an increased understanding of social presence will deepen the understanding of technologically mediated social experiences. Unlike the field of telepresence studies, therefore, it is not just concerned with virtual environments, but it also considers social experiences that are mediated by more inexpensive and popular means (such as digital communication technologies, specifically online chat technologies, that are

\(^{2}\) Teleoperation occurs when a machine (esp. robot) is operated at a distance.

\(^{3}\) Virtual environments may be defined as interactive simulations of real or imaginary worlds (see e.g. Carlsson & Hagsand 1993).

\(^{4}\) Some make a point of distinguishing between telepresence and virtual presence. Thomas Sheridan (1994:1) argues that “presence” has two manifest forms: “telepresence, wherein the human participant feels herself to be present at a location other than that which is actual (real and immediate), and virtual presence (or virtual reality or artificial reality), wherein the human participant feels herself to be present at a location which is synthetic, created only by a computer and various visual, auditory or haptic displays.”
the focus of this dissertation). Investigations into social presence appears in studies of subjects such as computer-mediated communication, virtual environments, artificial intelligence, human-computer interaction, online education, and online therapy, and that are conducted by scholars with backgrounds in a wide range of academic disciplines, such as psychology, philosophy, anthropology, education, informatics, cognitive science, media studies, and—more rarely—sociology.

Broadly speaking, one can say that the field runs two interconnected but distinct, lines of theoretical enquiry. The first one attempts to capture the essence of social presence by describing and defining—conceptualizing—what the phenomenon of being there together truly denotes. The second line of enquiry in the field has less to do with the essence of being there together, and more to do with how online social presence compares to offline social presence. Here it is often assumed that social presence always exists in its essential form in face-to-face interaction, and the task of the researcher is not to identify exactly what that essential form is, but to identify which elements of the face-to-face interaction are required for the phenomenon to occur, and which ones are not. Therefore, what is at issue in this branch is not the phenomenon of being together in itself, but the potential qualities and elements of social meetings that make the phenomenon possible.

As social presence is linked to several desired outcomes, advancing the knowledge of the nature, causes and effects of social presence is expected to have practical benefits. It is argued that if communication technologies can be engineered to enhance the occurrence of social presence, this will, among other things, improve the overall user experience, make it easier to establish and maintain relationships, enhance team efficiency, and make online education as well as medical and psychological online treatment programs more efficient and satisfying. In its applied form, the social presence construct is often used as a standard in which the qualities of digitally and physically mediated social experiences are compared. Here, an underlying assumption is that, if it is possible to copy the essential characteristics of physically mediated social presence, it will also be possible to design communication technologies that afford social experiences that are as alluring and rewarding to people as offline socializing (see e.g. Short et al. 1976; Zhao 2003).

Judging from the many scientific and practical advancements that social presence research hopes to contribute to, this is a very promising field. But not everyone shares this optimism. Fourteen years have passed since Swinth and Blascovich (2002:8) observed that “[a]fter nearly 30 years of theorizing and research, it appears that no one is clear about what social presence or co-presence are, let alone whether or not they contribute to our understanding of technology-mediated social interaction”, and the present situation is largely unchanged. In fact, despite its shortcomings, attempts at improving the theoretical foundations of social presence research dramatically decreased after having received much attention in the first half of the last decade, and in
recent years the concept is predominantly operationalized in empirical studies. But as there appears to be no general preference for one definition of social presence over the other, two empirical studies that claim they are both investigating ‘social presence’ might in fact be targeting quite different phenomena.

Another issue that has been raised by some is that, even though the field is specifically trying to understand social presence, the theories generally do not to any greater extent explore potential social, or interactional, sides of the phenomenon (Rettie 2003; Mennecke et al. 2011). Instead the field favours technological or psychological models of explanation. At the same time, a common assumption in the field is that social presence is unproblematic in face-to-face interaction. This leads us to the second field of research that I would like to introduce, which is the sociological study of the meeting between humans, a study area that usually falls within the domain of microsociology.

The Sociological Study of Social Meetings

Microsociologists study the bits and parts of social interaction (a concept that is ordinarily used synonymously with face-to-face interaction in this field): how actors behave, why they behave in certain ways, and what is collectively produced during a social meeting. Areas of study therefore include socialization, role dynamics, self-presentation, impression management, agency, interaction rituals and norms, and social emotions.

If there is one thing that most scholars in this field would probably agree about, it is that the label ‘unproblematic’ does not apply to face-to-face meetings. In fact, one could say that the study of human meetings is the study of the often-problematic collision between the individual and her freedom (or agency), and society as represented and reinforced by the participants of the meeting. It takes into account how the agency, intentions, and emotions of actors are reflected in their behaviour, while also considering how this behaviour is socially structured, for example through the behavioural norms and roles that are imposed on actors. Between those two poles, microsociologists often find a great deal of tension.

It is therefore regrettable—and perhaps due to a preference for psychological models of explanation—that the field of social presence research has not to any significant extent taken to heart microsociological observations on

---

5 Much microsociological theorizing has its roots in the philosophical traditions of American pragmatism (e.g., J. Dewey, C. H. Pierce, W. James, G. H. Mead) and/or phenomenology (e.g., E. Husserl, A. Schütz, J.-P. Sartre); however, there are also some exceptions, for example Georg Simmel’s sociological writings. Some notable perspectives in microsociology are symbolic interactionism (coined by Herbert Blumer, sometimes used to describe the works of C. H. Cooley and E. Goffman), social constructivism (P. L. Berger and T. Luckmann), phenomenological sociology (A. Schütz), and ethnomethodology (H. Garfinkel).
how actors are being together in face-to-face contexts and how these physically grounded meetings are not without their own problems. Erving Goffman (1967) has, for example, pointed out that face-to-face meetings are by no means protected from situations where participants for various reasons drift away and become ‘alienated’ from what is going on in the meeting, and Alfred Schütz (1932/1967) suggests that, even though the physical distance between actors may remain the same, the social distance between them can increase or decrease throughout the interaction. What could easily be seen as a problem with classical microsociological theory—namely that it takes the physical co-location of actors in interaction for granted—here becomes an asset, because it means scholars have been more focused on what is going on (and going wrong) in interaction, rather than on the mediation of what is going on (and going wrong).

At the same time, the focus on face-to-face interaction in microsociological theory remains problematic in the context of this dissertation, because the frameworks are not directly applicable to mediated interactions, and therefore they are not directly applicable to how people are being together thanks to digital communication technology either. In fact, many of the classical microsociologists would probably have argued that under no circumstances could mediated communication technologies ever allow people to be together. The focus on face-to-face interaction in microsociology can perhaps partly be explained by the fact that many of the frameworks and concepts that are in use today were created before the advent of the Internet. However, many of them were still developed after the invention of the telegraph and the telephone, yet these forms of mediated interaction were seen as “marginal and derived forms of social contact” (Goffman 1971:70) and were given a rather step-motherly treatment. So even if microsociological theories might provide some answers to what it means to be together, it is unlikely that they alone can satisfactorily answer the question of what togetherness entails in the digital era.

Nonetheless, there have been previous attempts at adapting microsociological concepts in studies of digitally mediated interaction. But few of these studies have specifically targeted what it means to be together in the digital world, and there is little in the way of a formal microsociological theory of digitally mediated interaction. Meyrowitz’ (1985) attempt at marrying Goffman’s interactionist approach with media-theoretical perspectives still stands as one of the more ambitious attempts at developing a broader microsociological theory of media and everyday life. However, Meyrowitz is predominantly concerned with mass media (like television), and less so with interactive media (like the telephone). Naturally, in No Sense of Place Meyrowitz also does not take into account the more recent developments in digitally mediated communication technologies.

The dissertation opened with a discussion of the phenomenon of being together, and has now come to a point where theoretical conceptualizations of
this phenomenon are in focus. The two discourses I have just introduced are each at the same time relevant to the overall topic of the thesis (what it means to be together in digitally mediated interaction), and problematic. Social presence theorists attempt to tackle the problem of what it means to be together in a technologically mediated somewhere, but, at present, their theoretical conceptualizations of the phenomenon leave much to be desired. By contrast, microsociology offers rich insights into everyday togetherness, and predominantly considers social or interactional aspects of how actors are being together. However, its theoretical models and concepts are developed for face-to-face interaction, and cannot be directly applied to mediated forms of social contact.

In the next section, I will continue to consider the example of online chat conversations, which was briefly introduced in the opening section and which represents the type of mediated interaction that the remainder of this thesis will predominantly focus on. The next section also further illustrates the need for a more interactional outlook on digitally mediated togetherness.

The Example of Online Chat Conversations

Digital communication technology offers a way to discuss the question of ‘disembodied’ or ‘distanced’ togetherness, not as an abstract possibility but as a concrete phenomenon. Out of the many options for digital communication that users can choose between, applications through which they engage in online chat stand out as particularly interesting. In its broadest use, the term ‘online chat’ can refer to any type of online conversation that happens near-synchronous or in real-time, meaning it can include not only text-based but also video-based conversation. Here I shall use the term more stringently, to refer only to text-based conversations. Some common alternatives for online chat are instant messaging clients such as WhatsApp and Google Hangouts, browser-based chat rooms such as Zobe, and Internet Relay Chat (IRC).

Online chat services mainly employ short text messages as the means of information transmission, but many applications, such as WhatsApp and Hangouts, also allow users to send small data files (containing, for example, photos, video- or audio clips) and they may also provide the opportunity to make phone or video calls. Unlike the instant messaging applications just mentioned, the online chat room Zobe is by contrast limited to the transmission of text only, and therefore does not allow phone calls, video calls, or file transfer. A more general difference between instant messaging clients and online chat rooms is that the former typically require users to be added to each other’s ‘contact lists’ prior to conversation, which makes users less anonymous to each other than they normally are in chat rooms. In addition, instant messaging often allows users to send messages to someone who is
not currently online (meaning the exchange will then resemble that of e-mail or short message services, SMS), while in chat rooms users can usually only send messages to people who are logged in to the service at the same time. In the following, the terms ‘instant messaging’, ‘online chat’, and ‘text-to-text interaction’ will be used to refer to the transmission of text messages that occurs in a situation in which all participants are simultaneously online.

While many digitally mediated technologies rely on the transmission of text, online chat differs from other forms of text-based communication technologies (such as the ones used for e-mail or online forum conversations) in that the messages sent are typically shorter, the message exchange frequency higher, and the time delay between conversational turns shorter. Even though each of the mentioned forms of text-based communication afford the same near-synchronous transmission\(^6\) of messages (that is, a sent message arrives with the receiver almost instantly in each case), users tend to treat online chat as a continuous conversation, thereby expecting themselves and other participants to be involved in, and responsive to, the conversation from its beginning to its end. In that sense, the conversational flow in online chat has more in common with a phone call, a video call, or even a face-to-face conversation, than it does with e-mail exchange or postings on online forums. It is a flow of interaction rather than an asynchronous exchange.

At the same time, text-based online chat differs from phone, video, or face-to-face conversations in that it does not allow users’ bodies to directly participate—users cannot hear, see, touch, or smell each other the way they can in face-to-face conversations (and in limited ways in phone and video calls). And so, while the structure of an online chat conversation is similar to that of more ‘embodied’ methods of communication, it is also different because it does not rely on any of the transmission methods used in face-to-face interaction—such as spoken utterances, body language, haptic cues, or olfactory cues. So in summary, one could describe online chat as a near-synchronous and disembodied form of communication that users typically treat as a synchronous conversation. Therefore, it is reasonable to suspect that online chatters can feel as if they all, at one and the same time, share a conversational “space,” yet also feel that their bodies are not present in this joint space. This makes the text-based interaction offered by online chat services an interesting case to discuss when considering what it means to be somewhere together, when the ‘somewhere’ is something other than a shared physical place.

---

\(^6\) Online chat applications may sometimes also support “real-time” transmission of text. That means that the receiver can see what the sender is writing while she is writing it, so that messages unfold on the screen letter by letter. While the sender can still revise her messages, the receiver will be able to see when a word is erased and replaced, a typo is corrected, or a sentence structure is altered. Among online chat users, however, the preference appears to be near-synchronous transmission over real-time transmission.
In the previous section I mentioned how, even in face-to-face meetings, actors could sometimes sense that they are still not precisely ‘there’ with each other. While there may be enough proof to establish that someone else is physically present with the actor, she may still apprehend that the other person (or she herself) is absent from the interaction or the conversation. I also described how my online chat conversation with Joan the Chatbot did not instil in me the same sense of being with another actor as the conversation with my (human) friend did. Interestingly, Alan Turing’s (1950) famous work on artificial intelligence can be used to further explore how this dynamic occurs in online chat conversations. The Turing Test (also called The Imitation Game) was designed to answer the question “Can Machines Think?” but one can argue that the test addresses the social capacities of machines, rather than their abilities in human-like thinking.

The game is played by one (human) interrogator and two or more competitors (of which at least one is a computer and at least one is a human being). After having conversed with all contestants, the interrogator is asked to determine which of his or her conversation partners are human and which are computers. If the computer-competitor manages to fool the interrogator into believing that it is a human being, it wins the game7. Naturally, the computer has a major disadvantage compared to human competitors in that it does not possess a human body. In Turing’s (1950:435) opinion this was an irrelevant issue, as he did not “wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane.” Contestants are therefore separated from the interrogator: with the help of computer terminals they answer questions in writing, while at the same time being visually cut off from the interrogator’s gaze. By letting the competition take place in an online chat setting, the ‘cheating’ body is thus presumably eliminated from the equation and complete anonymity is secured. Turing (1950:434p) argued that his problem therefore had “the advantage of drawing a fairly sharp line between the physical and the intellectual capacities of a man” and that “the conditions of [the] game make these [bodily] disabilities irrelevant.”8 In other words, the disembodied character of the test supposedly grants the machine a fair chance against humans.

Turing (1950:442) believed that, within 50 years, computers would be sophisticated enough to “play the imitation game so well that an average interrogator will not have more than a 70 per cent chance of making the right

---

7 In the original design of the imitation game, the computer had to imitate a female human being better than a male human being did in order to win. But this method, to my knowledge, is not utilized in modern day Turing Tests.
8 Stating that the inability to see the other’s physical body provides ‘complete anonymity’ is, of course, debatable. In the context of the Turing Test it merely means that the observer is unable to form a first impression of the other before interaction has begun.
9 This statement stands in sharp opposition to more recent claims from scholars dealing with embodied cognition (see e.g. Wilson 2002 and Anderson 2003).
identification after five minutes of questioning.” But 65 years later, no computer has yet succeeded in this task. Disembodied text-to-text conversation, although deprived of a range of non-verbal social cues, is still a much too complicated challenge for computers and their programmers. At least this is true in situations where the human participant is aware that her conversation partner may be a computer. Not surprisingly, computers have a better chance of deceiving people in ‘natural’ text-to-text interaction, yet it typically does not take long before the human participant realizes he is talking with a computer, or terminates the conversation out of boredom (see e.g. Payr 2010).

Nevertheless, the Turing Test is still attracting interest. To a small fraction of the artificial intelligence community, one of the most prestigious acknowledgements is to win the Loebner Prize Contest, a contest that prides itself as being “the first formal instantiation of a Turing Test”. A number of chatbots compete against a number of human contestants, and the chatbot that deceives the most interrogators—or, in the (usual) case that no judge is deceived, the one that is perceived to be most human-like—wins the title of “The Most Human Computer”. In recent years an interesting side award has also been handed out during the contest, namely the award for being “The Most Human Human” of that year. It goes to the human contestant that convinces the most interrogators that s/he is, indeed, a human being. In 2008, this title was granted to American author Brian Christian, who afterwards released a book about the preparatory measures he took to ace the test.

If Turing originally asked how someone could appear as thinking as possible, Christian (2010) investigates how one can appear as human as possible. Christian argues that the Turing Test is an excellent way of identifying ‘essential’ human features, such as being in the possession of a mind, a soul, and intentions. But apart from broadening the scope a little, Christian’s focus is similar to that of Turing. They both tackle the problem as if it was a “unidirectional” one, in the sense that it is seen as an issue that only depends on what features the entity that is supposed to appear or pass as a human has. So what Christian is asking himself is how he can manufacture a self-presentation that will work in his favour (i.e. make him pass as human) regardless of who the judge is or of any specific interactional circumstances of the online chat conversation.

This strategy, I would argue, overlooks one important detail: that social conversation is not a solo performance. Although the personal traits and other ‘fixed’ properties each contestant bring to the conversation are likely to influence its outcome, what emerges, dialectically, in the social interaction may be of far greater importance. For example, when judges are faced with narratives instead of being engaged in conversation they tend to have a much harder time determining whether a computer or a human is the originator of the text (Barber & Kudenko 2008). Therefore, one may want to question if human intelligence—or the most essentially human in humans—is what is
really measured in the imitation game, or if it rather targets how well ma-
chines engage in text-based social behaviour.

Likewise, it is improbable that the “Most Human Human” award is really
given to the human with the most ‘human’ qualities (an undertaking that
seems logically flawed to begin with). Any human that is mistaken for a
computer in the contest is, of course, still human—just as a highly scripted
and impersonal face-to-face conversation between a check-out clerk and a
customer is still an all-human exchange (even though a computer could po-
tentially also have managed to successfully follow such a conversational
script). Instead, it appears as if the award is given to the human that happens
to be taking part in the socially most efficient or rewarding conversations,
and it seems unlikely that this could ever be credited to the psychological
make-up or innate qualities of only one of the participants.

In summary, the disembodied yet conversational characteristics of online
chat make it particularly interesting to discuss in relation to what it means to
be (somewhere) together in modern day society. The failure of computers
(and sometimes humans) to pass the Turing Test further suggests that—just
as being in someone’s physical presence does not always involve an experi-
ence of being with that person—producing and sending text to each other
might not necessarily suffice for participants to experience that they are be-
ing together in online chat conversations either. Lastly, no matter how inter-
action is mediated, the question of what it means to be together in an online
chat conversation is probably a question that has more to do with what is
going on and being produced in the interaction, than it has with the specific
traits interactants bring to the conversation or with the medium in itself. In
other words, this is a question that is potentially better studied sociologically
than it is psychologically or technologically.

Purpose and Research Questions

The increased availability and use of digital communication technologies has
changed how we socialize with others on a day-to-day basis. In response to
this, the last 40 years have seen a growing academic interest in how digital
communication technologies affect our being somewhere together, or how it
possibly even creates entirely new forms of it, as it is believed that an im-
proved understanding of this phenomenon will lead to important scientific
and practical advancements. Yet, the very concept of being together remains
unclearly defined in the research that focuses on mediated interactions, and
the extent to which the field incorporates more interactional perspectives in
their theoretical models has thus far been limited. By contrast, microsocio-
logical theory studies how actors are together from more interactional stand-
points, but most of these theories cannot accommodate for mediated interac-
tions.
A large part of everyday social meetings are now mediated by electronic and digital communication technologies, and it is therefore an important task to ensure that our theoretical conceptualizations of micro-level social phenomena can accommodate for the fact that, overall, the mediation of many of those phenomena has undergone changes. As I have discussed, one area that needs improvement is the theoretical understanding of what it means to be somewhere together in a digitally mediated context, and the purpose of this thesis is to address this problem. More specifically, the thesis explores the problem of being together; examines what the concept of being somewhere together stands for in research on digitally mediated and face-to-face social contact; and attempts to identify, and offer some solutions to, the shortcomings in the overall understanding of modern day togetherness. While there are many different forms of digitally mediated interaction that exist at present, the example of text-to-text conversation mediated through online chat technologies has been chosen to serve as the primary case for theoretical analysis.

In line with the stated purpose, the following questions have guided the work:

1. Why has the advent of electronic and digital communication technology complicated the conceptualization of ‘being somewhere together’ (or ‘togetherness’), i.e. what is the underlying problem of togetherness?
2. Where are actors being together? Specifically, where can actors be said to be together when they are communicating through online chat technologies?
3. How is togetherness portrayed in social presence research?
4. How is togetherness portrayed in microsociological (specifically interactionist and phenomenological) research?
5. Given the underlying problem(s) of togetherness, how can the conceptualizations of togetherness be understood? How do the conceptualizations apply to interaction taking place in online chat arenas?
6. Which perspective is most promising for the future analysis of online chat togetherness, and why?

These questions also comprise an ambition to bring microsociological and media-theoretical approaches closer together. Although there have been several attempts at adapting microsociological theories and concepts in the analysis of digitally mediated social practice (e.g. Robinson 2007, Hogan 2010, Bullingham & Vasconcelos 2013, Beneito-Montagut 2015), most attempts at incorporating digitally mediated communication into general theories of society and the social have been made from a macrosociological perspective (e.g. Castells 1996). As noted by Rettie (2009), microsociology has thus far been comparatively conservative about allowing mediated interac-
tions into its general understanding of social interaction. However, the field can no longer approach mediated interactions as ‘exceptions’ that do not impact the everyday lives of individuals as much as face-to-face interactions do, because the reality is that the forming and maintenance of social bonds is increasingly being assisted by online chat applications and other digital communication technologies. It is my ambition, therefore, to make a contribution to the broader microsociological understanding of social life in our contemporary society, a society in which ‘digital’ socializing has become an integral component.

Methodology and Outline

The purpose and research questions stated in the previous section suggest that the appropriate way to conduct this study is through a theoretical investigation. By that I mean that the research purpose cannot be met by collecting and analysing empirical data of the phenomenon of ‘being together’ as it appears (or does not appear) in the everyday lives of everyday people, but that it rather calls for a theoretical investigation of how being together is discussed in the relevant fields of research. This does not mean that I will propose a new theory about being together. Instead, I study theoretical literature about or relating to togetherness in an attempt to better understand what researchers mean when they speak of togetherness. Throughout the thesis, the theoretical concepts and ideas concerning togetherness are discussed and analysed in relation to the example of online chat interaction. Based on my findings, I also suggest potential ways in which the theoretical understanding of togetherness in online chat arenas can be improved.

Empirical studies are often designed to be as transparent as possible when it comes to what data was used, and how it was collected and analysed. The overall quality of such a study is evaluated on the basis of the quality of the collected data and the methods used, and there are plenty of resources available for those who wish to improve the reliability of their research design. Theoretical studies, by comparison, are very different in that the concrete procedures a theoretician has undertaken are rarely explicitly described, and there is little in the way of formalized research methodologies to guide a theoretical investigation. That said, theoretical studies also make use of data, and of course they also result from some kind of method or procedure. However, the quality of a theoretical study is typically measured more on the basis of how strong and credible the argument is than it is on the data and methods used to arrive there (in fact, one could argue that the written argument is the method). In the remainder of this introductory chapter, I describe the outline of the dissertation, which should give the reader an idea of what material I used, how I approached the material, and what the constraints of the study are.
In Chapter 2 I begin my investigation by exploring the *where* of our being together, since surely the answer to what it means to be together relies on what type of *where* one thinks that actors are being together in. What *somewhere* do we have in mind when we say that we are with someone, *somewhere*? As far as ‘somewheres’ go, we might equally well be referring to a geographical location, our means of communication, or a social situation (but then what constitutes a social situation?). Of course, the purpose of the chapter is not to map all possible (concrete) places in which actors can be together, but rather to capture variations in the conceptualizations of, for example, space and place, or situations and environments, which can all be used to answer the question of where actors are. The material I use predominantly consists of microsociological and media-theoretical literature that deals with questions relating to environments and situations. The examined *wheres* are at the end of the chapter categorized into three types, and for each category I describe what the where of online chat would amount to.

Alongside this attempt to survey different *wheres* to be in, there is also a desire to better understand perhaps the most prominent reason why digital communication technology has spurred an interest in the phenomenon of being somewhere together, which is that it challenges our views on where we can be and socialize. In this respect, the chapter reads like a background to the theoretical problem of togetherness. At the same time, it also provides a frame for the coming chapters by suggesting that, depending on which *where* a scholar is focusing on, there are a number different ways in which the problem of being together can be approached and understood. In addition, the chapter sketches out the general interactive affordances of face-to-face environments and digitally mediated environments. Insofar as one can say that digital communication technologies create new *wheres* of interaction, this exercise highlights the core characteristics (and differences) between digitally mediated *wheres* and the *where* of face-to-face interaction. However, the section on digitally mediated environments also highlights the more specific characteristics of online chat environments.

In Chapter 3 I review what ‘being together’ stands for in two discourses. The first part examines theories of digitally mediated ‘togetherness’ as they are proposed in the field of social presence research. In the second part of the chapter I review how the microsociological literature has dealt with questions relating to togetherness in face-to-face interaction. In the summarizing discussion following each part I highlight different shortcomings in and aspects of the perspectives to be further examined in chapter 4.

Thereafter, in Chapter 4, I go on to discuss the reviewed literature on being somewhere together with the distinctions between different *wheres* of togetherness in mind, and I suggest that the views of two discourses on being together could be said to differ not primarily because they focus on two different modes of interaction, but because they focus on different *wheres*. So while the social presence literature can be said to target actors’ co-existence
in an environment, the microsociological literature discusses how actors share, or are ‘co-situated’ in, a social situation. I then continue to discuss the implications that these differing approaches have for the study of together-ness in online chat interaction, and I conclude that the situational perspective is a more promising direction for future research. In Chapter 5 I summarize the previous chapters, discuss my overall findings and the limitations of the study, and suggest some ways in which research on digitally mediated togetherness can move forward.

In summary, the thesis consists of five chapters, and each of the chapters answers some specific question(s). The two questions asked in Chapter 2 are what the underlying reason for the recent interest in the phenomenon of being somewhere together is, and how the where of being together can be understood. These questions might appear too disparate to fit into the same chapter, but in fact both questions concern the where of our being together. In Chapter 3, I ask how ‘being together’ has been conceptualized in social presence theory and in microsociological theory. Chapter 4 thereafter builds on the previous chapters in that I am seeking the answer to how, based on the findings from Chapter 2, the reviewed discourses in Chapter 3 can be understood. Here I also evaluate the two discourses and end by suggesting that the interactional approach is a more promising way forward. In Chapter 5 I briefly summarize the thesis and its overall conclusions, and I give some directions for future research.
2. Being *Where* Together?

The reason social presence—or ‘being there together’—has entered the scientific spotlight appears to primarily come down to the recent technological developments of electronic and digital communication media. One possible explanation for this is that social presence research is concerned with what it means to be *somewhere* together, at the same time as the use of electronic and digital communication technologies complicate the picture of this *somewhere*. As Meyrowitz (1985) observed in relation to the telephone, electronic communication technologies bring the difference between physical and social ‘location’ to the forefront. This difference started to become even clearer about a decade after Meyrowitz’ work was published, when computer-mediated communication, and in particular the World Wide Web, became accessible to the wider public and gradually changed the patterns of our everyday social interactions.

Certainly, on a philosophical level, the fact that the question of *where* we are cannot simply be answered by pointing to some geographical coordinates, is not a novel discovery (see e.g. Heidegger 1927). However, it is safe to assume that when social interaction only took place in shared physical locations, there was little reason to distinguish between social and physical ‘locatedness’ in the analysis of concrete social scenarios. But as our day-to-day interactions have shifted from being exclusively confined to face-to-face encounters, to being a blend of ‘online’ and ‘offline’ encounters, the difference between social and physical location has gained a concreteness that supersedes that of an abstract philosophical distinction. Up until the present day, the advancements of communication technologies have allowed our interactions to travel farther and farther away from the geographical location of our bodies, with the effect that the *where* of our being together is no longer a question only to be dealt with in hypothetical scenarios, but one that is becoming increasingly relevant in the analysis of rather concrete ones.

Drawing on sociological, phenomenological, and media-theoretical literature, this chapter examines how the *where* of our social encounters has previously been depicted, and considers some of the problems scholars have identified in these conceptualizations. The chapter is intended to serve as a backdrop to the coming chapters, in which the theoretical understandings of being together are reviewed and thereafter analysed. The coming sections give a background to the (theoretical) problem of togetherness, in that they detail how communication technology complicates the traditional under-
standing of where actors are, or must be, when they are being together. The many different *where*es in which the literature suggests that actors can be together are summarized in the last section into three broad categories, and I use the example of online chat to illustrate the implications of each category on the understanding of where actors are when they are using the technology. The categorization is later (in Chapter 4) employed to discuss the differences between how, in relation to online chat, togetherness would be described by social presence theory and microsociology, respectively.

**Environments, Situations, and Social Situations**

When we describe where we are, we typically refer to the environment surrounding us. But the ‘surrounding environment’ in this description is also sensitive to the context in which one is trying to define where one is. Whether a person says that she is in her kitchen, in her house, in London, in a phone call, in a relationship, or in the midst of writing a thesis, is entirely dependent on who is asking, and what is most relevant to the overall description. We are essentially referring to some aspect of our present circumstances that is currently the most meaningful. So in its everyday use, the surrounding environment can stand for anything from a geographical location, to a social or non-social activity, to a relationship, to a state of mind, and the list goes on. In essence, almost anything that we can say we are in could pass as an environment.

On a theoretical level we typically want to be more specific than that, for example by distinguishing between the environments surrounding us, and the situations we are in. According to Dreyfus (1991:146), Heidegger, for instance, made it clear that the ‘there’ of his Dasein (i.e. ‘being-there’), does not refer to a location in space, but to the “equipmental whole” that Dasein is currently most closely concerned with. Dasein, then, “brings its there with it. Dasein is its there” and for “each Dasein there is the situation as organized around its activity”. So for a particular Dasein, Dreyfus (1991:165) notes, Being-there is the same as Being-in-a-situation. Following this, then, whenever we are saying that we are ‘in’ something we are referring to our situation rather than an objectively defined environment. Therefore, one could perhaps also say that the difference between a situation and an environment is that the situation is dependent upon the beings that are in it, while the environment exists independently of beings, as ‘objective’ parts of the world.

In sociology, the ‘environment’ of an actor is, by contrast, sometimes described as all that which is within an actor’s perceptual reach at one particular point in time. Traditionally, the perceptive range of an actor has sometimes also been divided into the non-social elements in reach, and the social elements in reach. For example, Schütz (1932/1967:170, footnote) defines an environment as “[t]hat part of the *external world* which I can directly appre-
hend. This would include not only the physical but also the social environment with all of its cultural artefacts, language, [et cetera]” (my emphases). Defined in this way, the physical and social environments bear the characteristics of what Heidegger would call a situation (as it is dependent upon the actor that perceives it) while the ‘external world’ here appears to refer to the objective environment. However, the shapes and boundaries of the environments Schütz refers to are also firmly reliant on the objective world, as the objects of perception that Schütz mentions (i.e. physical materials and e.g. cultural artefacts) would supposedly exist and persist even if there was no one there to perceive them. And so the ‘environment’ could here potentially stand for the part of an actor’s situation that is most closely concerned with the ‘objective’ external world/environment. In any case, the difference between situation and environment is not so clear here.

In the following I examine the relationship between the physical environment (in the objective sense) and the social situation as it is described in traditional microsociological theory. Thereafter I outline what happens to the conceptualization of where actors are together when digital communication technologies enter into the picture. First I describe a problem with the traditional view on social situations that Meyrowitz has pointed out, and his alternative definition of the social situation. I then go on to describe how media has been conceptualized as environments, and some issues associated with this practice. On the one hand, it has been noted that the spatial metaphors that both scholars and technology users employ to describe one particular mediated environment, namely the World Wide Web, have brought with them their own issues that may in turn serve to further complicate the understanding of where we are together. On the other hand, thinking of media as environments lends itself to discussions of what these environments afford the actor, for example, whether they afford social situations. Drawing on Meyrowitz (1999), I propose that the notion of informational environments can be used to capture affordances (of both physical and digitally mediated environments) that are of relevance to actors’ being together. I then proceed to a brief discussion of the informational affordances of face-to-face environments and digitally mediated environments.

Physical Environments and Situations

To begin with, the physical (or spatial) environment is associated with geographically bound locations. It thereby becomes an area populated by material objects, artefacts, and living bodies, such as furniture, walls, DVD collections, human bodies, streets, buildings, cats, trees, and so on. Following this, the physical situation—or where I am in the physical environment—is defined by how geographically far my direct bodily senses can reach in this specific moment, and thereby my perceptual reach defines the boundaries of my situation.
The limitations of the body—most notably that it has to be located in a physical location to directly experience objects and events going on in that location; its inability to be in two physical locations at once; and the fact that it cannot travel in time—in certain ways structure an actor’s perceptual reach and, in effect, her situation. Right now, I cannot directly perceive the Eiffel Tower, because my body is in a house in London and my direct senses are not advanced enough to perceive (the physical materialization of) the Eiffel Tower from this location. I can, of course, experience it if my body is transported away from London to Paris, but as it will take me some time to get there I cannot perceive it in the present moment. So whenever we are situated in a physical location (which happens to be always), we are bound to a certain spatiotemporal structure that Berger and Luckmann have called the “here and now.”

The reality of everyday life is organized around the 'here' of my body and the 'now' of my present. This 'here and now' is the focus of my attention to the reality of everyday life. What is 'here and now' presented to me in everyday life is the realissimum of my consciousness. The reality of everyday life is not, however, exhausted by these immediate presences, but embraces phenomena that are not present 'here and now'. This means that I experience everyday life in terms of differing degrees of closeness and remoteness, both spatially and temporally. (Berger & Luckmann 1966:36)

Even though Berger and Luckmann here acknowledge that phenomena that are not currently present to me ‘here and now’ still make a difference to my understanding of my immediate reality, they maintain that the ‘here and now’ remains the most ‘real’ and significant zone of our everyday lives.

Closest to me is the zone of everyday life that is directly accessible to my bodily manipulation. This zone contains the world within my reach, the world in which I act so as to modify its reality, or the world in which I work. In this world of working my consciousness is dominated by the pragmatic motive, that is, my attention to this world is mainly determined by what I am doing, have done or plan to do in it. In this way it is my world par excellence. (Berger & Luckmann 1966:36)

From this statement we also learn that the ‘here and now’ situation is not only defined by the perceptual reach of my body, but also by whether or not I can directly manipulate it. A book can thereby be part of my ‘here and now’ in the sense that I can see it with my eyes and, by direct bodily manipulation, move it from the kitchen table to the shelf. But as I cannot directly change or manipulate the reality presented to me through the contents of the book, that reality is not part of my ‘here and now’. Similarly, the ‘here and now’ presupposes that, apart from just perceiving it, my body can also move in the situation and thereby in various ways manipulate it.
In summary, the physical environment is what is always around us, but whenever we say that we are being in something we are speaking of our situation. In fact, one can argue that we are never directly in an environment, but that we are rather in situations that are in turn held together by the spatial/material environment and its constraints. In the next section, I discuss what the literature has to say about actors who share a physically ‘anchored’ situation.

The Different Layers of the Social Situation

In sociology, the concept of a social situation tends to be used to describe something that supports, surrounds, or in various ways influences and structures, social interaction. That also means that the social situation is described as the situation where we engage in social interaction. There are, however, differing views on what the concept of a social situation denotes.

Symbolic interactionism, Gonoz (1977) observes, portrays each social situation as unique, new, and non-repeatable. This would imply that each situation is entirely created within the current interaction, or perhaps even that a situation is nothing more than the understanding actors have of the situation. But Goffman (1964:134) rejected this “happy-go-lucky” way of treating social situations, because it indicates that “social situations do not have properties and a structure of their own, but merely mark, as it were, the geometric intersection of actors making talk and actors bearing particular social attributes.” Instead he proposed the following definition:

I would define a social situation as an environment of mutual monitoring possibilities, anywhere within which an individual will find himself accessible to the naked senses of all others who are ‘present’ and similarly find them accessible to him. According to this definition, a social situation arises whenever two or more individuals find themselves in one another's immediate presence, and it lasts until the next-to-last person leaves. (1964:135)

This is a definition of the social situation that does not require a whole lot from actors: it is a state in which an actor is close enough to perceive or ‘access’ the information another person transmits, while at the same time knowing that the other also has the ability to access the information the actor himself sends out. But beyond this mutual awareness, actors do not have to be in any further state of reciprocity to one another.

Because of its emphasis on the actors’ ‘naked senses’ and the immediacy with which information is sent and received, the social situation here becomes inseparable from the ‘here and now’ situation described earlier—so much so that the social situation is, in fact, nothing more than a ‘here and now’ with two or more people in it. Perception and direct bodily manipulation also play key roles in the definition. As soon as the others are no longer
perceptible to the actor, the social situation vanishes, while the mutual accessibility between actors suggests that (even if they may not act on it) actors have the option, in different ways, to directly manipulate (or affect/influence) each other, for example through body language or speech.

But apart from defining how, at a particular moment, actors can perceive one another, the concept of the social situation can also include the various cultural and social rules and guidelines that shape actors’ behaviour. The social situation thus determines what behaviour is appropriate during a particular meeting, but also which actors to include in, and exclude from, the encounter. In this sense, the concept of a social situation refers to a particular system of behavioural norms. For example, the social situation of a funeral carries a range of behavioural rules and expectations that attendees will have to adapt to (or else face the risk of social sanctions). When an individual leaves the funeral to instead participate in the social situation of, say, a Christmas party, her behaviour and her role will change accordingly.

Drawing on Bateson, Goffman (1974:10) proposed the concept of frame to refer to a basic structure that organizes actors’ interpretations of what is going on, and in effect their behaviour. In that sense, the frame consists of both a meaning context and a behavioural structure that carry normative expectations of how actors should act in the social situation. Individual actors in individual meetings do not arbitrarily generate these principles of organization; rather, they arrive at an understanding of the frame and then attempt to adapt their behaviour in accordance with it (1974:247). While Goffman of course acknowledges that actors may also interpret or negotiate the behavioural structures imposed by frames differently—thereby acknowledging that actors too take part in the making of the frame—his stance is more structuralistic than the symbolic interactionist approach mentioned above, in that he disapproves of how the statement “[i]f men define situations as real, they are real in their consequences” (Thomas & Thomas 1928) has been taken to mean that the social actor is inherently free to define situations however she likes.

Elsewhere Goffman (1963:19) used the term ‘social occasion’ to refer to a wider organizing structure or social context in which social situations are likely to form, dissolve, and re-form and in which a more or less fixed pattern of conduct is recognized as the appropriate or ‘official’ one. The occasion is typically bounded in time and place, and accompanied by fixed equipment. It can also have a sharp and clear structure, or a rather diffuse one. For example, the funeral and the Christmas party mentioned above are social occasions with clearly defined structures, in that the course of these occasions is ordinarily outlined in advance, there are agendas for activity, and specific and clearly stated rules for appropriate conduct. More diffuse occasions (also called behaviour settings) carry similar, albeit less explicitly stated, structuring properties, yet they are not necessarily perceived by participants as entities having structures of their own, for example in the case of
a “Tuesday afternoon downtown”. While the idea of social occasions would be included in what Goffman (1974) later called a frame, it illustrates the rich variation in the types of frame that govern behaviour in social situations.

Goffman’s description of the frame is also intrinsically linked to the ‘here and now’ situation. The social occasions of the funeral and the Christmas party are, for example, spatially separate, so in order for the actor to move from one occasion to the next, she also needs to change her spatially defined situation. In addition, the physical environment not only influences perception, but it also provides props—like the coffin at the funeral or the decorated tree at the Christmas party—that in various ways participate in the social. Thereby, the concept of a social situation comes to stand not just for a shared ‘here and now’, but for everything at once: the perceptual reach I currently have of the bodies and objects in my immediate physical environment, my understanding of what is going on, and the norms and structures that influence and organize my behaviour (both in the wider—or more generalized—sense, and in the more specific sense).

Even before considering the additional complexity introduced by electronic and digital communication technologies, the question of where actors are together emerges as a rather complicated one. To summarize the discussion so far, there are, broadly speaking, three different wheres in which the actor could potentially be said to be with others. To say that actors are together in a physical environment would be to refer to the actors’ positions on some specific geographical coordinates, and the objective, material conditions that the environment imposes on the meeting. By contrast, to say that actors are in a social situation would be to point to how the spatiotemporal proximity between actors allows them to be mutually aware of, and accessible to, each other. Lastly, a framed social situation would be a where that considers how spatiotemporally proximate actors are also placed within a normative system (which in turn can comprise several different types of sub-frames). In the next section, I will examine what happens to the where of togetherness when the social meeting goes ‘online’.

Environments, Situations, and New Communication Technologies

In 1985, Meyrowitz pointed out that the focus on physical location in the then existing analyses on situational effects on behaviour was not surprising, because:

[unt]il recently, place-bound face-to-face interaction was the only means of gaining ‘direct’ access to the sights and sounds of each other’s behaviours. The physical barriers and boundaries marked by walls and fences as well as the passageways provided by doors and corridors directed the flow of people
and determined, to a large degree, the number, type, and size of face-to-face interactions. (Meyrowitz 1985:35)

The physically anchored situation was thus the only situation where actors could have direct perceptual and manipulative access to others.

But binding the conceptualizations of social situations to shared spatial situations becomes problematic when the conditions for social interaction change. Zhao (2006), notes that, as our communication technologies now allow us to interact in real-time at a distance, everyday reality is no longer confined solely to the zones of ‘here and now’ and ‘there and then’. Instead, Zhao argues, the reality of everyday life now also includes a ‘there and now’ environment that began to take shape with the invention of the telephone, but was not fully realized until the advent of the Internet. This new social zone resembles that of the ‘here and now’ in that its interactions take place in (near) real-time, but differs in that the bodies of participants may be located in vastly different spatial places. By increasing the immediacy by which messages can be communicated, new communication technologies have effectively reduced the importance of the spatial in our spatiotemporal locations. As the physical location does not structure the ‘there and now’ situation like it does the ‘here and now’, traditional conceptualizations of the social situation are difficult to apply to the new forms of social interaction that exist today.

But Meyrowitz (1985:36) made the observation that the existing understanding of the social situation was not, in its totality, useless for the study of mediated social life, even though the insistence on binding social situations to shared physical locations was misguided:

[I]s it actually [geographical] place that is a large determinant of behavior, or is it something else that has traditionally been tied to, and therefore confused with, place? [---] Indeed, a close examination of the dynamics of situations and behavior suggests that place itself is actually a sub-category of [the] more inclusive notion of a perceptual field. For while situations are usually defined in terms of who is in what location, the implicit issue is actually the types of behaviors that are available for other people’s scrutiny.

So although the geographically defined environment determines our perceptual reach, and thus also matters to the social situation, Meyrowitz suggested that there are other, more important, factors at play here.

For example, Goffman’s (1959) analysis of front and back regions is typically directly linked to the physical environment, like when the front region dining room of a restaurant is separated from the back region kitchen by a closed door. However, what really distinguish one region from the other are the different frames that govern the conduct of participants. If the physical outline of the restaurant was indeed the distinguishing factor, Meyrowitz (1985) argues, then the behaviour in the social situation of the kitchen would
remain the same even if a customer entered through the door. But, as Goffman (1959) also realized, all participants will start behaving in a front-region manner as soon as a customer (i.e. an “outsider”) enters this back region, which means that the frame has changed, even though the physical location remains the same.

Meyrowitz’ (1985:37) solution was to conceptualize the social situation as an information system, by which he meant a “pattern of access to information about others,” where ‘information’ stands for social information (i.e. knowledge about the behaviour and actions of ourselves and others). This conceptualization has the benefit of not only being applicable to face-to-face interaction, but also of allowing for the inclusion of interaction that does not take place face-to-face. Meyrowitz (1985:334) acknowledged that an information system could also be called a behavioural system, but preferred the former term because “‘information’ serves as a better common denominator to connect live encounters and mediated interactions”:

It is difficult initially to think of what we learn about an author through a book, for example, as access to ‘behavior,’ but it makes sense to consider it as access to a form of ‘information’ about the author and about the characters, places, and events described in the book.

To Meyrowitz, the informational content of a particular system is less interesting than the particular distribution of access to this information, that is, who gets access to what information, and who is in control of the information. For example, an actor’s inclusion or exclusion from a specific social situation, or information system, has less to do with his location in space and more to do with whether or not he is granted access to the information necessary for inclusion. Of course, in order to gain social information I still need to be able to perceive it, meaning that the access I have in an information system is the same as my perceptual reach.

The notion of information systems has been criticized for expanding (and thereby weakening) the concept of a social situation so much that it no longer serves its analytical purposes (Rettie 2009; Ito & Okabe 2005). For example, the idea of the social situation as an information system does not, like Goffman’s concepts of frames and situations do, take into account actors’ sense of sharing a social place, and the social expectations that the place in itself (through the cultural meaning ascribed to it) is imposing on them. The example cited above, where book reading is portrayed as a (socially) situated activity, further illustrates the width of Meyrowitz’ concept, where a social situation no longer needs to involve at least two actors who are mutually accessible to each other. However, what Meyrowitz’ concept accomplishes, is to show how perceptual (social) reach, a basic element of all social situations, is not determined solely by the physical location of actors, but also by
factors such as individual motivations to share information or the access to communication technology.

While Meyrowitz wanted to do away with the strict distinction between physically and electronically anchored social situations (by proposing that all social situations are patterns of information flow), he often employed metaphors that alluded to the spatial environment to describe the conditions of mediated meetings (for example by using the concept of middle-regions). In the next section, I describe how the use of spatial metaphors in the conceptualization of mediated *wheres* can sometimes be problematic, particularly since the advent of Internet technologies and the World Wide Web.

**The Non-Spatiality of the Digitally Mediated Environment**

When the Internet emerged as a medium for social interaction, a common approach was to describe it as creating entirely new arenas that existed in isolation from the physical environment (see e.g. Turkle 1996). In turn, the conceptualization of these ‘virtual’ arenas typically relied on spatial metaphors. Graham (1998:166) notes that, as ‘cyberspace’ is “no more than an abstract flows of electronic signals, coded as information, representation and exchange,” depicting it in terms of geographical metaphors is, perhaps, the most effective way of making sense of, and visualizing, this new environment.

Thus, an Internet point-of-presence becomes a web *site*. The ultimate convergent, broadband descendant of the Internet is labeled the information *superhighway*. A satellite node becomes a teleport. A bulletin board system becomes a virtual *community* or an electronic *neighborhood*. Web sites run by municipalities become virtual *cities* (see Graham and Aurigi, 1997). The whole society-wide process of technological innovation becomes a wild-west-like electronic *frontier* awaiting colonization. Those ‘exploring’ this frontier become web *surfers*, virtual *travellers*, or, to Bill Mitchell (1995:7), electronic *flâneurs* who ‘hang out on the network’. The Internet as a whole is variously considered to be an electronic *library*, a medium for electronic *mail* or a digital marketplace (Stefik, 1996). And Microsoft seductively ask ‘*Where* do you want to go today?’ (Graham 1998:166, emphases in the original)

The digitally supported social meeting thus gets linked to a ‘virtual’ location in which actors are ‘virtually’ co-located, just like the face-to-face situation has often been likened to the physical co-location of actors.

What is perhaps the main reason that the Internet, and particularly the World Wide Web, is so easy to describe in spatial metaphors is that the majority of Internet web ‘sites’ are persistent. They may also be accessed, and, often to some degree, modified by actors operating independently of one
another. In addition, user interfaces\(^\text{10}\) are usually deliberately designed to resemble spatial environments (Dourish 2004). The rationale behind this is entirely pragmatic: using spatial models when designing the user interface is thought to give users an intuitive understanding of how the system works. That is because the designer can assume that all users have in common a basic understanding of the spatiality of the physical environment, and so mimicking features of the physical environment is believed to improve the overall quality of the user experience (Dourish 2004:88).

Massively multiplayer online worlds (MMOWs) and massively multiplayer role-playing games (MMORPGs) take the forms of ‘virtual worlds’ and provide an obvious example of how spatial metaphors are incorporated in the design of user interfaces. Virtual worlds are persistent computer-simulated environments that a user accesses and inhabits through her ‘avatar’ (i.e. a graphical representation of the user). According to Pearce and Artemesia (2009:20p), a virtual world is a two- or three-dimensional spatial construct that is geographically contiguous, meaning it can be mapped and explored. The user can move around within the world with her avatar and to varying degrees manipulate the environment. The world is also persistent in the sense that it is always ‘there’ online, and users’ actions in it are cumulative and consequential. One user’s activities thus become a part of the world, and of other users’ experiences of it. Interaction between users typically occurs in real-time and is predominantly channelled through text, avatar gesture and audio. All these characteristics serve to, as closely as technologically possible, mimic the structure of the physical world as we know it. The more the limits of the virtual environment resemble the limits of physical environments, it seems, the better\(^\text{11}\).

Online chat platforms and instant messaging applications are in comparison less intricately and deliberately designed to mimic spatial environments. Yet in order to engage in online chat or instant messaging, a user must enter a chat room or open an instant messaging window, and while doing so she will be represented by an online appearance (such as, for example, a profile picture or a user name). Similar to the interaction in MMOWs and

---

\(^{10}\) In simplified terms, the user interface is what allows for interaction between a human and a machine, such as the computer. The computer ‘desktop’ is thus a user interface because it allows the user to give the computer commands and observe how the computer responds to these commands. The part of a website that is visible to a user is called the web user interface.

\(^{11}\) But, on a parenthetical note, the striving for spatial realism does not extend to the virtual body inhabiting these worlds. In the MMOW Second Life, for example, users’ avatars do not need to eat or sleep and will never die or get injured. They can also at any time change into new shapes, and fly. So while the design of the virtual world appears to follow some ideal of familiarity and realism, how the virtual body can appear and act in the environment gives the impression that it is limited only by imagination. Realistically, the appearance and possible behaviours of the virtual body are, of course, restricted by the technology and by decisions made by the programmer, but the portrayal of a body that transcends the physical body is still noticeable.
MMORPGs the interaction in online chat is thus also described as located in places, which users populate by means of their ‘virtual’ bodies. A user must ‘be’ in a chat room or an instant messaging thread in order to modify what is happening, and some software (especially online chat rooms) also require the users to be ‘there’ in order to not miss out on what is going on. If the user’s computer runs out of power she will probably miss out on everything that is happening in the chat room while she is logged off. Instant messaging applications, by contrast, typically store information so that the user can be ‘away’ from the conversation and still take part in it, responding to what everyone else has texted at a later point. In that sense, chat rooms share more in common with a physical room than instant messaging ‘rooms’ do, in that one cannot expect to first-handedly take part of the information being shared in a chat room if one isn’t ‘in’ the room as it happens.

As I have shown, online applications are often easy to think of in terms of spatial environments. The same is not true for many other media. A book, for example, is persistent—I can put a book I have just read on the shelf and leave it without supervision for years and years, knowing it will continue to exist. But in those years, nothing will have changed in my book—the author will not have been there to rewrite a chapter. So even though the environment described in the book has durability, neither I, nor anyone else, can participate in it once it is published. As a medium, the book only allows us to be passive consumers, not active participants. Conversely, a telephone call supports joint participation in the creation of its content, but not durability. It exists only so long as at least two people participate, and when that joint participation ceases, the ‘environment’ of the phone call in itself vanishes. I cannot hang up on my friend and then expect her, and the conversation, to still be ‘there’ in my phone the next morning. Of course I can record a phone call and listen to it afterwards, but that recording will then have the same character as the book does—it supports duration, but not participation.

To return to the digitally mediated environment, a consequence of the use of spatial metaphors to conceptualize it is that we often think and talk about the ‘online world’ as an entirely new environment, completely separate from the physically anchored one, in the same way that different physical environments are spatially separate. The ‘virtual’ environment is thus put in binary opposition to material (physical) environments, where the latter is often also described as more ‘real’ than the former. This is, for example, reflected in how Internet users have come to adopt the term “In Real Life” to refer to things and events happening outside of the mediated environment, as if what is happening ‘online’ somehow does not belong to reality. This idea is not limited to everyday discourse, but is also reflected in the academic literature. In the following quote from Turkle’s (2011:153) pessimistic portrayal of modern day social contact, the idea that digital social life is not entirely ‘real’ prevails:
When part of your life is lived in virtual places—it can be Second Life, a computer game, a social networking site—a vexed relationship develops between what is true and what is ‘true here,’ true in simulation. In games where we expect to play an avatar, we end up being ourselves in the most revealing ways; on social networking sites such as Facebook, we think we will be presenting ourselves, but our profile ends up as somebody else—often the fantasy of who we want to be. Distinctions blur.

While researchers generally seem to have become more cautious about using the real/non-real, or virtual/non-virtual distinctions in their studies of digital social life, the coming chapter’s review of social presence theory illustrates that these distinctions are still employed in some more recent works.

In summary, spatial metaphors have paradoxically enough served to describe cyberspace as a non-spatial environment, which, according to Graham (1998:167), obfuscates “the complex relations between new communications and information technologies and space, place, and society:”

In the simple, binary allegations that new technologies help us to access a new ‘electronic space’ or ‘place’, which somehow parallels the lived material spaces of human territoriality, little conscious thought is put to thinking conceptually about how new information technologies actually relate to the spaces and places bound up with human territorial life. Without a thorough and critical consideration of space and place, and how new information technologies relate to, and are embedded in them, reflections on cyberspace, and the economic, social and cultural dynamics of the shift to growing ‘telemediation’, seem likely to be reductionist, deterministic, oversimplistic and stale.

In line with Actor Network Theory (ANT), Graham goes on to suggest that only by conceptualizing space, place and technology as relational and forever linked and co-evolving can we fully understand the inter-relationships between them, because “defining space and place separately from technological networks soon becomes as impossible as defining technological networks separately from space and place” (1998:181).

**The Digitally Mediated Environment as an Embedded Environment**

ANT developed as a response to the realization that social theory cannot (or will not) properly account for how non-social objects take part in the social world. Instead, it strives to maintain a sharp divide between the ‘social’ and ‘nature’. Latour (2005) suggests that one underlying cause of this shortcoming is a long standing wish to distinguish sociology from other disciplines, which spurred an ambition to protect the identity of social phenomena by showing how they were strictly separated from, for example, geographical, economical, legal, or psychological ones. As soon as this narrow definition of society and the social was established, it came to be taken as an a priori explanation for any sort of social phenomenon.
Once this domain had been defined, no matter how vaguely, it could then be used to shed some light on specifically social phenomena—the social could explain the social—and to provide a certain type of explanation for what the other domains could not account for—an appeal to ‘social factors’ could explain the ‘social aspects’ of non-social phenomena. (Latour 2005:3)

Because of this, John Law (1991:8) notes that sociology now lacks a method to “juggle both the technical and the social” because, being characterized as ‘non-social’, technology has not been “productively integrated in the sociological imagination.” ANT holds that our reality is continuously constructed and reconstructed through an endless amount of ongoing interactions, both between the living, and between the living and the non-living. Moreover, it is here maintained that the non-social influences the social just as much as the social influences the non-social. The social, therefore, does not determine the non-social any more than the non-social determines the social. Similarly, technology is being shaped by society just as much as it is actively shaping society. In this sense, ANT harmonizes with Gibson’s (1986) ecological view of environments.

Gibson developed his theory in an attempt to capture how “the possibilities of the environment and the way of life of the animal go together inseparably” (1986:143), and of particular interest for this study is his concept of affordances. An affordance is neither an objective, nor a subjective, property of the environment, but points both to the environment and the observer: the restraints or possibilities for action that the environment imposes on actors are at the same time both objective and fixed, and relative to the actors. While media, substances, surfaces, objects, places, and other beings in the environment afford one actor certain behaviours and not others, it may afford completely different behaviours to a second actor. Affordances do not merely present themselves to actors, but have to be actively perceived as affordances in order to influence the actor’s behaviour. Therefore, affordances belong both to environments and situations: every environment has a finite and fixed number of possible affordances that exist independent of actors, and every situation offers a selection (if not all) of those affordances. But the number and nature of this selection is in turn dependent on the perceiving actor12.

12 For example, an open field of grass affords stable ground to stand on for both a cow and a lion, but it only affords nourishment to the cow. If an actor walks into the field while the cow is eating the grass, the cow becomes part of the environment/situation and will, as such, afford things to the actor. If the actor is a lion, the environment/situation will now afford a prey. If the actor is a human being, the environment/situation will afford a prey or fresh milk. But to neither the cow, the lion or the human being do the environment/situation afford shelter. Humans can, to some extent, alter the affordances of their environment/situation, but—and here is the twist—only as long as the environment/situation affords the opportunity to perform such appropriating acts. The human can, for instance, build a house for shelter on the grass field, but only if the environment/situation affords material for the house.
Human beings also tend to alter—or appropriate (Ingold 1987)—their environments in order to change what these environments/situations can afford. In the beginning of human history, the physical environment only afforded one type of arena for social contact: the face-to-face situation. After thousands and thousands of years of appropriating the environment to create new affordances, the human being has now arrived at a point where she can appropriate her environment to also afford other arenas for social interaction, such as the digitally mediated interactive arenas that are of relevance to this thesis.

That does not mean that in all physical environments, at all times, is this affordance of an additional arena for communication present, nor does it mean that, when it is afforded, people always choose to act on this possibility of extended perceptual reach. People still engage in face-to-face interaction—sometimes because they have no other choice, and at other times because face-to-face interaction has more to offer than digitally mediated interaction does. That is because digitally mediated interactive arenas come with their own affordances, which in turn limit the possibilities for action that actors have, such as in what way this arena can be appropriated. Because digitally mediated arenas offer specific ‘affordance structures’, it still makes sense to talk about these arenas as environments, only they must be conceived as embedded in, and themselves afforded by, the physical environment, rather than independent and separate from it.

So the environment shapes the social in that the environment provides us with affordances, but the social also shapes the environment in that humans’ actions in the environment are both selective (even if the environment affords the opportunity of action A, the actor can still choose not to engage in activity A) and consequential. My actions in an environment will alter the environment, sometimes to the extent that it affords the ability to carry out a new type of action. From this follows that the digitally mediated social situation always also depends on, and in some way incorporates, physical locations and physically located social situations, just as the physical world of today cannot be separated from the digitally mediated one. The world within our reach—the ‘here and now’—now also includes the ‘there and now’.

It is not difficult to think of scenarios exemplifying this point. A person who commutes to work may use her laptop to send e-mails to her colleagues while travelling, a teenager who is having an after-school milkshake with his friends may be on his smartphone sending instant messages to another friend, a wife may spend her Wednesday evening sitting on the couch next to her husband, debating what colour the dress13 is on an online social network-

---

13 “The Great Dress Debate” refers to the worldwide confusion that occurred in February 2015 after a member of the social media website Tumblr posted a photo of a dress and asked what colour other members of the website thought the dress was. As the background lightning in the photo allegedly made different people perceive the colours of the dress differently (i.e. either as black and blue or white and gold), the photo rapidly spread across a multitude of
ing website. In all of these examples, the person using communication technology is at the same time situated in a physical environment and (in these cases) also in physical proximity to other social beings, and it is likely that this person’s digital communication activities are coloured by his or her physical surroundings, and vice versa. The commuter’s e-mail communication might be interrupted by another passenger’s inconsiderate behaviour, another teenager at the coffee shop might hijack the first teenager’s instant message conversation, and the husband might also involve himself in The Great Dress Debate that is unfolding on his wife’s computer screen. Even if neither of these latter things happen, the fact still remains that our digital communication practices are already taking part in the shaping of face-to-face situations and interactions. Likewise, going ‘online’ does not mark a sudden and definite separation from the material world, not only because digital communication technology is made possible and supported by material bits and pieces and bodies, but also because the ‘offline’ world in various ways permeates digitally mediated conversations and discourses. As Latour (2005) would have it, the material world is always already a part of the social, and the social a part of the material, just as the ‘virtual’ is a part of the ‘physical’ and vice versa.

The discussion leading up to this point would suggest that the question of where we are when we are together simply cannot be answered by referring to the binary opposition of virtual and material environments, because there is no such binary opposition. But that raises the question: can we no longer make any distinction between physical environments and digitally mediated environments, or between face-to-face situations and digitally mediated ones? A simple answer would be that of course we can, and probably should, employ such a distinction, for the same reasons that it is useful to analytically distinguish between face-to-face situations that take place in different physical environments. But as Graham (1998) and actor network theorists have pointed out, focusing too much on the spatiality versus non-spatiality, or materiality versus virtuality of environments and situations might interfere with the aim of better understanding the (singular) reality of contemporary social life. That being said, we must not forget that, at least when it comes to digitally mediated environments, actors themselves are also actively defining these environments as (virtually) spatial places that stand in opposition to ‘Real Life’ places, and it is, perhaps, reasonable to assume that this influences actors’ own subjective understandings of where they are together.

ANT and the ecological perspective help us see the digitally mediated environment as an ‘embedded environment,’ but it does not give this environment more specific contours. Next, I therefore wish to turn to Meyrowitz’ later conceptualization of the medium as an informational environment, as online social platforms and gained substantial attention throughout the world. See also Rogers (2015).
this can be read as a specification and concretization of the medium as an embedded, informational environment.

The Medium as an Informational Environment

‘Medium’ is, as we have seen, an elusive concept and it can be used in a variety of ways. Sometimes, Meyrowitz (1999) observes, the word ‘medium’ gets to represent a rather innocent vessel that either holds or sends content without changing or manipulating it. For example, no matter if I tell a story about what happened today in an instant message (i.e. in writing) or over the phone (i.e. in spoken language), I will think of it as the same story, because what I consider to be the significant parts of the story can be delivered through both media. At other times, the medium stands for a type of language where each medium has its “unique range of expressive potential” (1999:46). Translating a story told through the medium of the telephone to another medium, like instant messaging, will therefore always modify the story in itself. For example, it is unlikely that I would use the same words and phrasings when telling a story through instant messaging as I would use when telling it over the telephone, and I would also have to use different strategies to nuance what I was saying (e.g. by using emoticons instead of changing my tone of voice).

A third way of using the word ‘medium’ is in terms of an environment, that—regardless of variations in content or ‘language’—has unique characteristics that distinguish it from other media. Environmental features, to Meyrowitz (1999:49), include the type of sensory information the medium can transmit; the immediacy of communication; whether it supports unidirectional, bi-directional or multidirectional communication; whether the interaction is simultaneous or sequential; and how easy it is to use and learn to use. An online chat environment would thus be defined by the available means for expression, the number of people that can communicate with each other at the same time, the time it takes for a sent message to reach its receiver, and so on. Even though he himself does not acknowledge it, what Meyrowitz is pointing to when he speaks of media as environments, is how media latently afford the user the opportunity to gain access to social information, for example by specifying the fixed and unique number of possible social actions that they offer.

This conceptualization is broad enough to be applied even to physical environments in which actors are situated, but it is also specific enough to only take into account the affordances that matter for actors’ co-situatedness in environments. The notion of mediated environments as informational environments furthermore appears to go hand in hand with Meyrowitz’ (1985) earlier idea of social situations as patterns of information flow, as the informational environment takes part in the shaping of such a pattern. As pointed out above, however, the concept of social situations as patterns of information flow blurs the potentially important distinctions between different
types of elements that shape the flow of information (e.g. between environmental affordances and frames). By contrast, the concept of the medium as an informational environment does not make any claims to cover the entire range of factors that determine a social situation (particularly not the framed ones), but covers only the environmental possibilities and constraints (i.e. the latent affordances) on those social situations.

As a last stop before concluding this chapter, I will briefly describe the face-to-face environment and the digitally mediated environment in terms of informational environments. In other words, I will sketch out and to some extent compare the informational affordances of these two ‘types’ of environments. The main reason for this endeavour is that a large portion of the literature reviewed in Chapter 3 tends to emphasize the distinction between face-to-face environments and digitally mediated environments, and therefore it appears appropriate to address the differences between the two here.

Informational Affordances of Physical and Digitally Mediated Environments

From the discussion so far one can conclude that every social meeting and situation is dependent on a supporting environment and that the transformation of environments that carry these supporting qualities has led to changes in social life. The environments of interest for this thesis are primarily digitally mediated ones (particularly those supporting online chat) and to a lesser extent physical ones. Since the two ‘types’ of environments (i.e. digitally mediated and physical) will be frequently discussed in the coming chapters, I will here give a broad description of each of them to clarify what—apart from the obvious—distinguishes the one environment from the other. When I discuss digitally mediated environments, I will oscillate between characteristics general to all of these environments and characteristics that are specific to online chat environments.

However, some initial clarifications are in order. Firstly, I will discuss the environments as informational environments, as this will capture the (physically and digitally mediated) environmental affordances that are most relevant to the social meeting. Without doubt, both environments can afford opportunities of varying importance to humans—the opportunities of shelter, entertainment or shopping, to name a few—and most of these affordances probably matter to the social meeting in one way or the other. In this case, however, the affordances of interest are those that more directly influence the flow and perception of social information, and those that define what social behaviour an actor can engage in in each of the two environments. Moreover, as what follows is a comparison between two types of environments, I am particularly interested in those affordances that can be said to distinguish the one environment from the other. For example, a given physi-
cal environment might be as good as a given digital environment at affording the opportunity to communicate in Swedish, but the shared physical environment is typically better equipped to afford spoken Swedish than the shared digital environment is. Meyrowitz’ (1999) list of features that distinguishes between different media environments can be used to guide the discussion (even though Meyrowitz might not have approved of the depiction of the physical environment as a medium environment). These features predominantly concern how the actor can transmit and access social information in the environment (i.e. possibilities for action and perception), and how the environment discriminates between different actors (e.g. who is granted access, how many actors can simultaneously participate, and so on).

Secondly, I will sharply separate the physical and the digitally mediated environment here. Digitally mediated environments are of course both made possible by, and integrated with, the physical environment. As has been discussed, digital communication technology often permeates our social meetings in physical environments, for example when a person is writing an instant message to a friend on her phone, while at the same time (seemingly?) listening to her partner who is sitting across the table and busy describing what happened at work. Similarly, digitally mediated environments are never isolated from physical environments. Any actor who participates in digitally mediated interaction is simultaneously tied to her physical environment, even though her current physical location might of course be devoid of other people (i.e. it is not a shared physical situation).

Nonetheless, I will be focusing here on what can be called ‘pure’ informational environments, where in the one case only the affordances offered to actors who are only situated in a physical environment are considered, and in the other case only the affordances that are offered to the actors by a digital environment are considered. The main reason for this is that the purpose of this exercise is to highlight and compare the differences between the two environments, and not to discuss the often-blurred borders between them.

**General Affordances of ‘Pure’ Physical Environments**

Whenever two physical bodies are located in close proximity, the physical environment affords actors the ability to be mutually aware of each other and to interact face-to-face. Face-to-face interaction is ordinarily the first type of interaction we learn to engage with others through, and we do not need any other ‘equipment’ besides our bodies to interact in this way. As Goffman (1963:15) would have it, we are therefore experiencing each other with our naked senses in face-to-face interaction. The access to face-to-face interaction is moreover shaped by the limitations of our bodies, for example we must be close enough to another actor to be able to directly perceive and be perceived by her.

The information transmitted by the physical body comes in different forms: cues can either be visual, auditory, tactile, or olfactory in nature. Be-
cause of that, we can say that shared physical environments afford the ‘richest’ possible interaction, as the information is ‘packaged’ in a variety of ways and received and perceived through several of our senses.

Goffman (1961, 1963) distinguishes between four types of social information, of which three are afforded in ‘pure’ physical environments. These four types are: expressed and communicated information (1961); and disembodied and embodied information (1963). The difference between communicated and expressed communication (or expressions ‘given’ and ‘given off’ as Goffman (1959) had earlier described it) comes down to intention. Communicated information is information that is given intentionally and purposefully, while expressed information represents the kind of information that interactants give off unwillingly and sometimes also unwittingly. Merely entering a physical environment with one’s body produces signs and marks—i.e. exudes information—and there is virtually no way an actor can be in such an environment without also expressing information. Shared physical environments also afford expressed information to be transmitted independently of communicated information: whenever two physical bodies are interacting, they are automatically transmitting information that the actors may not have intended to convey or are even aware of conveying. Expressed information is therefore not an action in its own right, writes Goffman (1961:6), but a by-product of actions. While an actor can (and often tries to) consciously act so as to manage what information she expresses, she cannot directly control expressions (because they then become communications).

The second distinction introduced by Goffman is the one between disembodied and embodied information. Goffman (1963:14) suggests that embodied information is the kind of information that an actor expresses or communicates by means of his own current bodily activity, and this information can only be received by the other so long as the bodily activity is sustained. Disembodied information is, by contrast, received in the absence of bodies (for instance in the exchange of letters) where the information is ‘trapped’ in something that holds the information long after the actor has ceased to inform. Therefore, disembodied information is not afforded in ‘pure’ physical environments.

Embodied information may contain elements of both expressed and communicated information, since it concerns the transmission of both largely ungovernable and inevitably ‘expressed’ information as well as more intentional ‘communicated’ information (that is, an actor can usually control some of the information transmitted by her body, but never all of it). But embodied information is also—under normal conditions—reciprocal, as the act of perceiving the information another body transmits in most cases also
involves letting one’s own (perceiving) body reveal information.\textsuperscript{14} Perception and transmission thus happen at the same time, which is what scholars mean when they talk about ‘synchronous’ or ‘immediate’ interaction.

To Schütz (1932/1967:163), immediacy covers both the amount and richness of social cues that the environment can afford (i.e. spatial immediacy), and the time delay between my observation of the other, and the other’s observation of me (i.e. temporal immediacy). In spatial immediacy, actors share a physical space with others and are fully aware of doing so. Spatial immediacy allows actors to use all perceptive channels to read others, at the same time as each actor—willingly or unwillingly—transmits a constant stream of bodily signals for others to apprehend. Temporal immediacy refers to a situation where the other person’s experiences are “flowing side by side with mine” so that “I can at any moment look over and grasp his thoughts as they come into being”. To share temporal and spatial immediacy, which is (currently) only properly afforded in shared physical environments, means that expressions are constant (i.e. both my body and that of the other always expresses something) and immediately observable, which in effect means that an actor can at any time check whether her interpretive scheme appears to correspond to the other’s expressive scheme.

As a physical environment is not just comprised of human bodies that are always within range for mutual perception, actors may have to take measures for it to afford face-to-face interaction. They must position their bodies to ensure that they can perceive (and be perceived by) others, and adapt to disturbances that may temporarily disable one or several perceptive senses (as for instance a very noisy environment can do). Also, they may need to alter the environment in ways that facilitate interaction (e.g. removing objects that block visual perception; turning down the volume of the television). These adjustments further serve to optimize what Schütz (1932/1967:168) refers to as a shared point of view.

One could say that our point of view falls under the category of spatial immediacy, as the angle from which we orient ourselves to one another will affect how many perceptive channels are in use, and how well we are able to receive transmitted cues. For example, if you and I are standing on separate sides of a crowded room, the point of view I have of you (that is, peripheral and predominantly reliant on my visual and auditory perception—perceptions that are, in addition, impaired by the distance I have to you and by people blocking my view and creating background noise), is damaging an otherwise immediate transmission of social cues. So even though the environment may afford a constant and varied stream of expressed social cues, the angle from which I perceive these expressions may still compromise the immediacy and richness of what I perceive. My point of view in effect influ-

---

\textsuperscript{14} Exceptions to this include different varieties of surveillance, where person X can observe person Y, while Y cannot at the same time observe X.
ences the conclusions I can draw about what is currently affecting you, for example what you from your perspective can see, hear, feel, and potentially be distracted by.

The physical proximity of actors; abundance of intended and unintended social signals sent and received; the instant transmission of both intended and unintended social signals; and its accessibility are all qualities that serve to justify why face-to-face interaction is generally, and perhaps rightfully so, considered to be the most efficient and genuine form of social interaction. Thanks to the affordances of spatial immediacy, temporal immediacy, and shared point of view, the physical environment has also been assumed to be the only environment that affords the constitution of a common intersubjective world (Schütz 1932/1967:171). Not only because the immediacy of the encounter makes us “grow older together” and enables participation in the “joint flow” of shared experiences (Schütz & Luckmann 1973:64), but also because, in Schütz’ (1932/1967:170) view, only individuals that share an undivided environment can “assume with more or less certainty that the table [they] see is identical in all its perspective variations to [others]”. To Schütz, then, a common intersubjective world is afforded when interaction is spatio-temporally immediate, because it means that the actor will experience that he has direct access to the totality of information that the other transmits, and also that the environment that he perceives is roughly the same as that of the other.

Digital Mediated Environments and their Affordances

Digital communication technologies mainly recycle old strategies of information presentation (e.g. text, images, audio, video), but they have revolutionized how fast information can travel over a geographical distance. In addition, the versatility of the technology allows individuals to employ a multitude of information channels during one and the same interaction session. Thus, digital interaction may be almost synchronous, i.e. ‘near-synchronous’, and compared to other media it also affords ‘richer’ information. As an effect, digital communication media share a greater resemblance to the face-to-face medium than earlier communication media could ever aspire to.

While digitally mediated environments afford various different media for communication, they all have in common that they do not afford physical, concrete objects (most notably physical bodies) to enter the environment (other than in a represented way). Naturally, digitally mediated environments do not offer a complete disconnection from the physical body but rather serve to extend its reach, as—even in digital interaction—our material bodies ultimately lie behind everything we project into the environment. Nonetheless, digitally mediated interaction exclusively relies on visual and auditory perception, since the environment cannot afford the transmission of tactile and olfactory information. In that sense, it is less ‘rich’ in information than
face-to-face interaction is. While the structure of online chat resembles that of face-to-face conversation (in that the message exchange rate is high and users tend to perceive an online chat conversation as having a coherent structure with a clear beginning and end), there is a significant lack of non-verbal information in these environments.

However, even though information is transmitted through fewer channels in digital interaction, communicated, expressed, disembodied (and sometimes also embodied) information is still normally afforded. That information is communicated and disembodied might not come as a surprise, but Goffman (1961) also observed that although expressed information is richest in shared physical environments, it is not exclusively confined to face-to-face interaction. Where there is communicated information there is always also expressed information as the manner in which language is used always simultaneously gives off at least some bits of expressed information. For example, in online chat, the choice of words, emoticons and punctuation marks can give an indication of the sender’s age and even nationality or educational background—even if the sender is not explicitly mentioning it. In addition, emoticons are frequently used to express the tone with which something is said, where a message ending with a smiling emoticon will be perceived in a very different way than one ending with a frowning emoticon. In that sense, the service affords the user an opportunity to reduce some of the uncertainty that comes with a lack of non-verbal expressions, but it may also reveal information about the user without her being consciously aware of it.

When digitally mediated environments afford actors the ability to directly speak to each other—either through a real-time video call where they can also see each other’s physical bodies while talking, or in a more traditional audio-only call—the environment also affords embodied information to be transmitted. However, the affordance of embodied information is still limited compared to the physical environment, as the only embodied cues that can be perceived by the other are the audible and, for video calls, visible ones.

Digitally mediated interactive environments do not afford the simultaneous transmission, reception and interpretation of cues. Although some types of digital interaction are more ‘temporally immediate’ than others (compare e-mail to a real-time video conversation), no digital environment can at present afford complete synchronicity, and thus it will also afford less temporal immediacy than physical environments do. Even the reciprocity of embodied information in an online video call is compromised in the digitally mediated environment, for instance because most video conferencing systems do not afford reciprocal eye contact.\textsuperscript{15} Compared to video calls, the reliance on text

\textsuperscript{15} There are various ‘gaze-correcting’ techniques available, but most of them are expensive and impractical. In the hope of reaching the average user of video conferencing systems, Kuster and colleagues (2012) have, however, discussed a method that is supposed to be inex-
in online chat of course makes it even less synchronous as it takes the sender some time to produce a message before he or she can send it to the receiver. Unless the service affords real-time text, the user will also receive little (if any) additional information about what the other is doing in between turns.

Since digitally mediated environments are not spatial environments, they naturally do not offer spatial immediacy. However, as was illustrated earlier with the example of virtual worlds (such as MMOWs or MMORPGs), digitally mediated interactive environments are often made to recreate a sense of spatial immediacy. Even if they are less intrinsically designed to mimic spatial environments than virtual worlds are, online chat environments also afford a type of durable structure that appears in a similar way to all users. Insofar as one actor can make references to the common structure of the online chat environment and expect the other to also be able to perceive these referenced things in a similar way, actors can in fact share a point of view in online chat. This happens in approximately the same way that actors can make references to what they perceive in a shared physical environment. In WhatsApp, for example, one user can instruct another to click on a certain button in the upper left corner of his or her screen (so that a certain feature of the application is activated), and this is of course only possible if both users perceive the interactive environment in a way that for all intents and purposes is the same.

However, as actors that share a digitally mediated environment typically do not share the same physical environment, the point of view they can share remains limited. An actor might be able to grasp the other’s point of view insofar as it concerns her perspective of the digitally mediated environment she shares with the other, yet she may simultaneously understand very little of the other’s ‘physically’ anchored point of view. However, video and phone conversations open up a small window into the physical environment of the other, where the actor may be able to see parts of the room the other is video-calling from and where she can hear sounds in the other’s background. But the grasp an actor can get of the other’s complete point of view in digitally mediated interaction is nonetheless very limited compared to what she can achieve in a physical environment.

Lastly, to access a digitally mediated environment, the individual at the very least needs a device such as a computer or a smartphone, network access, and some basic skills in using and communicating through the technology. So while everyone can (and must) access the physical environment, pensive, efficient and easy to set up. But as of 2016, reciprocal eye contact is generally not afforded in digital video interaction.

As was previously mentioned, in real-time text the receiver can see the sender’s message unfold on her screen, letter by letter, while the message is being created. This creates a sense of immediacy to the conversation, and it also allows the receiver to see when the sender is writing or correcting typos, replacing words, or changing the sentence structure similar to how a speaker may correct herself in spoken conversation.

---

16 As was previously mentioned, in real-time text the receiver can see the sender’s message unfold on her screen, letter by letter, while the message is being created. This creates a sense of immediacy to the conversation, and it also allows the receiver to see when the sender is writing or correcting typos, replacing words, or changing the sentence structure similar to how a speaker may correct herself in spoken conversation.
there are several roadblocks that have to be crossed in order to access the digitally mediated one. A person’s income, level of literacy, ability to use and learn to use the technology, and the Internet infrastructure in the area where she lives may all prohibit her from entering or successfully using the digitally mediated environment. Many online chat services are however free of charge (apart from the cost of Internet data) and relatively easy to learn. While some require the user to have a smart phone (e.g. WhatsApp), others can be accessed either from a smart phone, a computer, or a tablet computer (e.g. Google Hangouts and Zobe) meaning they are accessible to a broader range of people.

In summary, the differences mentioned so far between ‘pure’ physical and digitally mediated environments concern the number of perceptual channels supported, how fast information can travel between sender and receiver, to what degree actors can make use of the environment as a shared point of reference, and the social reach that actors are afforded. The physical environment affords ‘richer’ and more immediate interaction, and allows for greater control over other participants’ points of view. However, the social reach in face-to-face environments is limited compared to digitally mediated environments, as actors’ bodies must be located in close proximity to each other in order for interaction to be possible. By contrast, digitally mediated environments, such as online chat environments, afford actors the opportunity to interact in near immediacy despite being located in different geographical places, but they are inferior to physical environments in terms of immediacy and richness of information flow, and of how actors can understand one another’s point of view.

Discussion and Summary

As was stated in the introduction to this chapter, the emergence of social presence research at the end of the last century was not so much an effect of a sudden and unprovoked interest in the nature of ‘being together,’ generally, but of a growing realization that the situations that we are together in are no longer (and perhaps never were) identical to the environments in which our bodies might be physically co-located. In this chapter, therefore, I describe how, on the one hand, the relationship between the social situation and the environment, and, on the other hand, the relationship between physical environments and digitally mediated environments, have been portrayed in sociological and media-theoretical discourses.

When it comes to the first relationship, one way of distinguishing between an environment and a situation is to make a subject-object distinction. An environment then becomes a limited amount of ‘stuff’ that is possible to perceive, combined with a particular set of limits to, and possibilities for, action. It is ‘objective’ in the sense that an actor does not need to perceive or
act in the environment for it to continue being an environment. By contrast, a situation can only be realized by actors—the situation is defined by the actor’s perceptual range and the actions she understands that the environment affords her. For each action the individual undertakes in the environment she is situated in, the environment will be affected in some way, and so the relationship between situation and environment is mutually influential, as is also emphasized by Gibson’s (1986) notion of affordances. The distinction between environments and situations is not always employed in theories that centre around the behaviour or experience of actors, as, for instance, the phenomenological or microsociological ones referred to previously do. Here, the only environments of interest are those in which actors are situated, and so the distinction between environments and situations becomes rather blurred.

Social situations have traditionally been closely tied to the shared physical situation, or to the perceived co-location of actors’ physical bodies. But the social situation is also seen as everything which, in less tangible ways, structures actors’ behaviour. Added to the notion of the social situation is therefore also the behavioural norms imposed on interacting actors, and how actors interpret and negotiate these structures. This means that not only has the concept of the social situation been used to denote the physical environment surrounding actors, but it has also included the behavioural norms and expectations that structure their behaviour. Goffman would say that a social situation occurs as soon as physically co-located actors are mutually accessible to one another, while the behaviour co-located actors engage in occurs within a frame that guides it. This frame can, in turn, be seen to have (at least) two layers: the first layer being the more generalized normative structure to behaviour (which would include occasions), and the second involving specific structures of conduct negotiated within each particular meeting.

However, Meyrowitz made an important contribution to the understanding of behaviour in framed social situations by showing how the strong emphasis on the shared physical situation is misleading. In fact, when it comes to the framed social situation, the only thing that matters is what behaviour is possible and impossible, encouraged and suppressed. In that sense, the physical environment does matter, but only insofar as it shapes behaviour—its role in situations is therefore the same as the role of cultural and social norms and expectations. By defining the framed social situation in terms of a much less physically grounded, and therefore a much more inclusive, information system, Meyrowitz ended up with a conceptualization that he thought could be useful in the analysis of all kinds of social behaviour, including mediated behaviour. However, others have argued that the concept of an information system fails to capture the influence of place, norms, and expectations on behaviour (Rettie 2009), and the way in which Meyrowitz’ concept considers both the reading of a book and a live conversation between
two actors as social situations makes it too inclusive for the purposes of this thesis.

From this point I continue to discuss the problem of employing spatial metaphors in the conceptualization of media environments. This is particularly relevant in the context of the World Wide Web, because this digitally mediated environment is so frequently represented in terms of a ‘virtual’ spatial environment that is contrasted to the ‘non-virtual’ spatial environment. Not only has this resulted in a tendency to see the one environment as less real or important than the other, but it also gives the impression that, since we cannot be situated in two different physical environments at once, then perhaps we cannot be ‘genuinely’ situated in both a physical and a digitally mediated environment at once either. In line with ANT, Graham therefore calls for a more relational view, in which the digitally mediated environment is portrayed for what it is, namely an integrated extension to our physical environments, or a form of embedded environment.

By introducing Gibson’s (1986) notion of affordances, I continue to describe how media environments are ultimately afforded by—and within—physical environments. I then briefly review Meyrowitz’ (1999) idea of media environments as informational environments, where each medium is portrayed to come with a set amount of latent possibilities for information flow. Lastly, I discuss the general characteristics—or affordances—of physical and digitally mediated (informational) environments.

Given Graham’s (1998) call for a less polarizing discussion about the ‘virtual’ and the ‘non-virtual,’ this choice might seem odd, but this is a case of going where the literature reviewed in Chapter 3 points. Firstly because Chapter 3 consists of two parts of which the first one reviews literature that deals with togetherness in mediated environments and the other explores literature about togetherness in physical environments. Secondly, the distinction between digital ‘togetherness’ and face-to-face ‘togetherness’ is often made in social presence theory, which is why I deemed it appropriate to draw out the differences between physical and digitally mediated environments before moving on to review this literature. The discussion revolved around the affordances of immediacy, informational richness, shared point of view, and social reach offered by the two environments, where face-to-face environments afford more immediate and richer interaction, with better opportunities of sharing a point of view than digitally mediated environments do. However, digitally mediated environments allow people to engage in interaction even when they are not in each other’s spatiotemporal proximity, which means that these environments afford a better social reach than face-to-face environments do. Online chat environments typically also afford greater immediacy than many other text-based environments (e.g. e-mail, or online forums) but lower immediacy and richness than online services that are offering e.g. phone- or video calls.
So what does this leave us with when it comes to the where of our being somewhere together? One answer to this would be that the where is a social situation. But as I have intended to illustrate in this chapter, we might mean many different things by the term ‘social situation’. The social situation may stand for the sharing of a specific pattern of access to information; the awareness of other actors in an environment; or the social frame guiding co-located actor’s behaviour, both in the wider and in the more specified sense. By the term ‘social situation’, we might also refer only to the physical proximity of two or more actors, or to all variations combined. When media are involved, the mediated social situation can be considered as opposed to, and less real than, face-to-face situations—in a sense the mediated one can be represented as an entirely new species of social situation or frame—or the social situation can be seen as an information system whose overall structure remains the same regardless of what medium of communication is in use. The following is a tentative categorization of different wheres in which actors could be said to be situated together, both generally and as applied to online chat:

(1) Actors can be situated in the same informational environment. This would involve being exposed to a shared set of affordances of information flow, but actors do not need to have any established relationship of reciprocity going on between them. In order for this category to accommodate both digitally mediated environments and physical environments, it must also include instances where the shared part of an actor’s informational environment is not identical to the totality of the informational environment she is situated in, i.e. the complete and undivided informational environment of one actor must not fully correspond to the complete and undivided informational environment of another. In line with what was previously mentioned, the shared informational environment of online chat would involve the means of expressing and receiving the information that the application offers to all its users. It would also include actual information that users have shared access to, such as the appearance of the application interface and user created content (e.g. profile information and lingering messages).

(2) Actors who already share an informational environment can be together in a larger social situation. This would involve being currently in one another’s reciprocal, social reach, which assumes there is some form of mutual access and awareness between actors. This where draws its structure both from the informational environment and from more general frames that organize social behaviour. In the context of online chat applications, there are several ways in which reciprocal social reach can be established. In a chat room, the user can normally see who else is ‘there’ in the room and accessible for social contact by looking at the sidebar where logged in users’ names will appear. Whenever someone logs off, his or her name will disappear from the list. For instant messaging, on the other hand, the user typically has access to a list of her contacts, and the application will indicate who of
these contacts are presently logged in. Sometimes the application will also inform the user of when another user was last active, which will give an indication of whether his or her absence is temporary or permanent. Of course, for both types of online chat the user’s own availability will simultaneously be communicated to other parties. When it comes to the general social frames that organize behaviour in online chat, these would include the norms of appearance (e.g. what user name/profile picture is appropriate), norms regarding what language of communication or topics of conversation are preferable, and norms regarding who can contact whom and under what circumstances. This can be likened to how the behaviour in a public place is still socially organized, even when actors are not overtly interacting with each other. Actors normally follow some rules of appearance and conduct specific to the place (compare behaviour in e.g. an underground station to that of a bar), and they also have a sense of whom they can approach, and of when and how it is appropriate to approach them.

(3) Actors who share a social situation can also be together in a specified social situation. This would involve actors being reciprocally and overtly oriented toward, or even overtly responsive to, each other. The structure of the specified social situation is decided by the informational environment, the larger social situation, and by more specified normative structures or frames. So when a user of an online chat service is in a specified social situation it means she is engaged in interaction with one or several other users. Her behaviour will not only be structured by the informational environment and the general interactional norms previously listed, but also by the structures that emerge in this particular sequence of interaction. These structures are therefore affected by what each individual brings to the situation (e.g. personality, socioeconomic background, computer proficiency, writing skills, and so on) and also by factors such as the time of day or participants’ previous history of interaction.

I believe these categories are broad enough to encompass all the variations described in the chapter, and that each of them can also be used to refer to both physically mediated events and—as illustrated by the example of online chat—digitally mediated ones. However, it is important to emphasize that while one can analytically isolate the different categories from each other, in reality they overlap. Being in a larger social situation for example presupposes the sharing of an informational environment, and being in a specified social situation presupposes both the sharing of an informational environment and the sharing of a larger social situation. The categorization should thus only be seen as an analytical tool that helps in identifying different layers of the where in which social meetings take place.

Moving forward, the coming chapter will review what being somewhere together means in two discourses—social presence theory and microsociology.
This chapter examines what being somewhere together stands for in two discourses. The first discourse—social presence theory—is concerned with togetherness in digitally mediated environments, and the second discourse—microsociology—is concerned with togetherness in physical environments. While social presence theory is entirely focused on coming to grips with the phenomenon of being somewhere together (or social presence), in microsociology there is no similar area of study that explicitly targets the more elementary aspects of togetherness. However, several interactionist or phenomenological perspectives discuss social meetings from a microscopic viewpoint, and the examination of togetherness in the microsociological literature is therefore concentrated on these branches of the field.

The chapter consists of two parts, beginning with a review of social presence theory, followed by an examination of togetherness in microsociological theory. Following each part is a summarizing discussion that also comments on the implication of the various perspectives on online chat. The chapter concludes with a summary and discussion of both fields.

Social Presence Theory and ‘Being There Together’

In Chapter 2, I proposed that the underlying reason for the recent interest in what it means to be somewhere together is that, thanks to technological advancements, the social situation is no longer restricted to the shared physical environment. Therefore, it is no wonder that of the two dominating discourses in social presence theory, one is concerned with how the digitally mediated environment affords the transmission and reception of social information, and the other is concerned with the subjective reception of the available information in these new types of environment. However, an additional branch of social presence theory frames the phenomenon in more active terms. Here, social presence can be said to stand for a form of self-presentation, or the ability to successfully represent oneself in a digitally mediated situation. Most of these theories focus on how actors can successfully participate in educational settings in digitally mediated environments.

In the first part of this review of social presence theory, I will take a closer look at how social presence has been defined in terms of an affordance. There are two variations of this definition, namely social presence as a prop-
property of the medium and social presence as observability. In different ways, they both relate to what social cues can be transmitted in/by the environment. Thereafter, I review perspectives in which social presence is related to self-presentation. I then continue to review literature that frames social presence as the actor’s experience of the outside world. There are several differing takes on this. While some scholars see social presence as the experience of the medium, others see it as an experience of the other, and yet others see social presence as an ‘illusory’ experience in which the individual fails to notice that she is currently in a mediated situation. Here I also review the interactional elements that have been portrayed as influential to the experience of social presence. Last, I summarize and discuss the different understandings of the phenomenon of social presence.

**Social Presence as an Environmental Affordance**

The three perspectives on social presence examined here are all concerned with what types of social stimuli (i.e. social cues or social information) the environment is able to provide actors with. Although there are several variations of these ‘environmental’ definitions (I will discuss the most prominent ones in a moment), they all share the idea that social presence equals the amount, the quality, or the availability (or some combination) of certain types of social cue.

In the first perspective, social presence is portrayed as a property of the medium, or, more precisely, as the medium’s ability to convey social cues. In many ways, this notion is closely related to the second perspective, in which it is argued that social presence is the ‘observability’ of self and others in a digitally mediated environment, i.e. it denotes the extent to which actors are embodied and thus convey observable social cues.

**Social Presence as a Quality of the Medium**

Of the theories that consider social presence to be a question of medium quality, most belong to a family of CMC theories that is often called the ‘cues-filtered-out perspective’. Several of the earliest attempts to create a comprehensive theory of electronically mediated social presence—such as the media richness theory (Daft & Lengel 1984; Daft et al. 1987, Trevino et al. 1987), early social presence theory (Short et al. 1976) and the theory of media appropriateness (Rice 1993)—are standard examples of cues-filtered-out approaches. In essence, these theories take the medium itself as a starting point, and try to assess to what degree the medium is able to emulate face-to-face interaction (which is assumed to be the optimal mode of human interaction because it is superior in conveying social information, or social cues[^17]).

[^17]: Suggested examples of cues that can serve as social stimulus are: motion (Biocca & Harms 2002), non-verbal communication (Schroeder 2006), behavioural cues, such as "para-
The more ‘cues’ a medium filters out, the less effective/intimate/warm the communication will turn out to be. Hence, a medium capable of conveying a rich amount of social cues possesses more social presence than media that allow less perceptual channels to be used and thereby ‘filter’ out much of the (particularly expressed) information.

Even when social presence is defined as an experience or as self-presentation (as described below), the amount and richness of the social cues a medium affords is often considered highly influential on the quality of social presence (e.g. Biocca & Harms 2002; Schroeder 2006, Zhao 2003; Short et al. 1976; Kehrwald 2010a; 2010b; Heeter 2003; Mennecke et al. 2011; 2010; Swinth & Blascovich 2002). Similarly, the more synchronous or immediate the afforded means of communication is, the better it is thought to be for social presence (Swinth & Blascovich 2002).

**Social Presence as Observability**

Social presence as observability covers theories where social presence is described as the existence of social cues in a situation where someone is also capable of observing those cues. The social cues of importance for social presence may either be transmitted by the other, as the “social cues that signify the presence of others within some interactional context” (e.g. Swinth & Blascovich 2002:10), or by both subject and object (e.g. Heeter 1992; Schroeder 2006; Ijsstelsteijn, van Baren, van Lanen 2003; Mennecke et al. 2010).

Some scholars do not make any demands on “observational reciprocity” (i.e. that the observed person simultaneously observes his observer) (e.g. Swinth and Blascovich 2002), whereas others argue that it is essential that the observer at the very least believes that the other is also capable of observing him or her in return (Heeter 1992:2). Social presence thus becomes “the extent to which other beings (living or synthetic) also exist in the world and appear to react to you.”

Observability may not instantly come across as an environmental issue, as it is actors who are the sources of social cues. However, as all (human) actors have physical bodies, and thus the ‘ability’ to be observable, it is ultimately the environmental conditions of a meeting that can make observability problematic. As soon as we start considering environmental factors that alter perception, we also start seeing variation in observability. For example, two actors that are co-located in a completely dark room, where they do not touch each other and both remain dead silent, would not, according to the definitions presented here, be socially present. But if someone turned on the

---

linguistically generated” non-verbal signals and elements (Zhao 2003:451p); embodiment (Mennecke et al. 2011:425, see also Zhao 2003), stimuli representing the virtual environment and objects in that environment, cues contained in messages (Kehrwald 2010a:45), and the responsiveness of the other agent (Schroeder 2006).
light, they suddenly would be. In these examples, the cue-transmitting bodies of both actors remain the same, while the environmental conditions change. And it is that environmental difference that makes social presence (as observability) a reality in the one case, but not the other. The same principle naturally holds for digitally mediated environments too, as it is not the information transmitting capacity of actors’ physical bodies that make digitally mediated interaction vary in informational ‘richness,’ but the affordances of the media environment.

That observability is an environmental affordance perhaps becomes clearer when observability is discussed in the literature, not as social presence in itself, but as a necessary condition for it. Here, the affordances of embodiment and co-location, which both concern the observability of actors, are often held as necessary for social presence to occur (e.g. Zhao 2003:450, Kehrwald 2010a:40; Biocca & Harms 2002; Mennecke et al. 2011; 2010). Co-location is normally used to refer to circumstances where two or more users occupy a shared virtual space by means of their digital representations (see e.g. Kehrwald 2010a; Biocca, Harms & Burgoon 2003; Schroeder 2006; Zhao 2003), while a digital representation (also called an embodied representation) can refer to the most minimalistic digital representation of the user as well as to highly complex, three-dimensional avatars that the user can operate in a human-like way in a digitally mediated environment (e.g. in a virtual world like Second Life). As a general rule, the more realistic the embodied representation is, the better it is thought to be for social presence (Swinth & Blascovich 2002; Klimmt & Vorderer 2003; Daft & Lengel 1984; Mennecke et al. 2010).

Social Presence as Self-Presentation

Social presence has also been described in more active terms, as a form of self-presentation. In fact, Rettie (2005) argues that the emphasis on ‘expressive’ cues in the cues-filtered-out approaches mentioned above already hints at an understanding of social presence as self-projection, but she continues to point out that these perspectives do not explain how self-presentation is actively performed and jointly constructed in interaction.

The complexity of presentation of self challenges the simplicity of social presence, which is treated as a straightforward construct, sent and received either directly in face-to-face interaction or with some loss through mediated channels; the elements of joint construction and collaboration in interaction is ignored. (Rettie 2005:358)

Social presence as self-presentation, she suggests, needs to be seen as a social activity shaped not only by the technology, but also by actors.
Within the field of online education research, social presence is similarly framed as a type of ability to present an attractive self or the activity of presenting oneself in the digital classroom. Social presence is, for example, described as the learner’s ability to “project” him- or herself into a so-called community of inquiry (Rourke et al. 2001, Garrison 2007; Kehrwald 2010a), an ability to communicate successfully despite unfamiliar social conditions or limitations for interaction (Bostan 2009; Kehrwald 2010b:33), or an ability to establish relationships (Garrison 2007). Partly on the grounds that social presence is an ability to self-project, Kehrwald (2010a:47) proposes that social presence is a form of subjective agency where “[i]ndividuals make choices about what sort [of] and how much relational information to convey. In doing so, they retain privileged access to their own subjective mental states whilst granting limited access to the others with whom they interact.” The individual thus becomes socially present within the interactional context by projecting communicative cues and making herself visible (Kehrwald 2010b:36). Elsewhere, Kehrwald (2010a:46) portrays the social cues “demonstrating” social presence as more or less advanced and intentional self-descriptions. For instance, communicating one’s own subjective perspective, which includes carefully selecting what aspects of one’s immediate point of view and prior experiences one should convey, would be one way of creating social presence.

Therefore, in this branch of social presence research, the emphasis on self-presentation and management of self-presentation is strong. While Rettie highlights the collaborative process behind self-presentations, the educational perspective is more concerned with social presence as a form of social competence that will make it easier for actors (in this case students) to successfully partake in the digitally mediated social situation (here an online classroom).

Social Presence as an Experience of External Events

In the field of social presence research, the most common way to approach the phenomenon of social presence is in terms of a psychological one. Even so, there is a wide array of suggestions as to exactly what type of internal phenomenon the concept denotes\(^{18}\), but the vast majority of scholars do agree

that the phenomenon in some way relates to how actors experience the outside world.

So when social presence is defined as a subjective experience it is usually described as a response to objects and/or events external to the subject. Included in this category are not only passive, non-reflective responses but also reactions associated with a certain degree of intrapersonal agency and conscious will, such as judgments. Although there is no general agreement on precisely what external event or object the subjective response ought to be linked to, most researchers tend to define social presence as an internal response to other social beings. Finally, in one of the earliest conceptualizations of social presence, the phenomenon is described as a response to the medium in itself. A common idea in this category is that social presence is a form of illusory experience of what is going on in the mediated environment, where the actor is led to believe that what she experiences is, in fact, ‘unmediated’.

**Social Presence as an Experience of the Medium**

The following perspectives are ‘media’-centred in that they define social presence in terms of how the medium is experienced. In some instances, the nature of an actor’s experience of the medium tends to be quite vaguely described. For example, apart from defining presence as a property of the medium or as a response to another being, Short and colleagues (1976:66pp) also suggest that social presence denotes the actor’s judgment of the medium. Many scholars draw on Short and colleagues’ media-judgment definition, for instance Tu (2000:28), who defines social presence as a “subjective quality of the medium [...] which depends upon the medium’s objective quality”, or Schroeder (2002a:3), who suggests that the concept denotes “a measure of the individual’s psychological state in relation to the medium.” However, it is more common in this branch to see social presence as a type of ‘non-experience’ of the medium, which I will discuss next.

**Social Presence as the Experience of High Social Realism or as Distorted Perception**

Some scholars propose that social presence is the experience of high social realism, where face-to-face interaction is considered the most ‘realistic’ form of social experience (see e.g. Biocca, Harms & Gregg 2001:3, Rice 1993, Zhao 2003, Kumar & Bensabat 2002; Rettie 2003). In essence, social presence is here understood as the extent to which users experience digitally mediated interaction in more or less the same way they experience face-to-face interaction, and the more closely digitally mediated interaction comes to resemble face-to-face interaction (i.e. the more ‘socially realistic’ the medium is) the higher the likelihood that this experience will occur.

A similar line of thought is that social presence is an illusion that occurs when the user experiences mediated social events as ‘real’, or ‘non-
mediated’ (see e.g. Mennecke et al. 2011, 2010; Klimmt & Vorderer 2003:346; Lee 2004; Heeter 2003). Again, it is here assumed that the more digital interaction resembles face-to-face interaction, the greater the likelihood that the user will be ‘struck’ by social presence, or the ‘illusion’ of being with others. In fact, the highest levels of social presence can, in this view, only emerge if the medium is completely unnoticed as “a psychological state in which technology users do not notice the para-authenticity of mediated humans and/or the artificiality of simulated nonhuman social actors” (Lee 2004:499, see also Zhao 2003:451).

Social presence as an illusion occurs when an observer observes technologically mediated social cues as if technological media did not transmit them. In other words, the observer will apprehend technologically mediated social cues as if they were mediated face-to-face. For instance, Gunawardena and Zittle (1997:9) write that social presence is “the degree to which a person is perceived as a real person in mediated communication,” while Lee (2004:499) asserts that social presence is a state in which “technology users do not notice the para-authenticity of mediated humans and/or the artificiality of simulated nonhuman social actors.” Furthermore, Ijsselsteijn and Riva (2003) assert that one aspect of social presence is the “disappearance of mediation”. This disappearing mediation can, according to the authors, also be described as a phenomenal unawareness of both the technological medium and the physical environment.

These descriptions of social presence can therefore be summarized in terms of people’s private experiences or judgments of what (via the medium) the physical or virtual environment affords. If the descriptions involve a response to the social beings inhabiting—or the social interaction that occurs in—these environments, the medium’s capacity to emulate physical actors and face-to-face interaction is still what determines whether there is going to

---

19 The term ‘non-mediated’ is of course a little unfortunate as one can argue that, in face-to-face interaction, the body serves as a medium (and thus face-to-face interaction is also mediated).

20 The notion of presence as an illusion was first proposed by Lombard and Ditton (1997), who claimed that all existing accounts of (tele-)presence (including those notions of telepresence that have a more social character) share the common denominator “perceptual illusion of non-mediation” (1997:9).

21 Others have higher demands on the “illusion”. Social presence is not merely to have the illusion of observing another social being, in this view. Instead, social presence is to have the “illusion of being with a mediated person” (Klimmt & Vorderer 2003:346), or, in Heeter’s (2003:341) words: “social presence implies that being with someone virtually feels like being with them physically.” Klimmt and Vorderer (2003:346) further maintain that the illusion of being with another does not sufficiently describe co-presence (which, according to the authors, differs from social presence). In co-presence, the subject has the social presence-illusion of being with a mediated person, but added to that is also the impression of “being present in a mediated room.” Even though they do not mention the very concept of “illusion”, Mennecke and colleagues (2011:425) assert that embodied social presence, too, consists of the cognitive confusion of digital interaction with “real world BtB [i.e. body to body] interaction.”
be a social presence response or not. Moreover, social presence tends to be seen as a given fact in face-to-face environments (i.e. all shared physical situations are assumed to automatically afford the social presence response in humans), meaning the concept is only useful in relation to digitally mediated environments. Here, what is non-mediated is also understood as ‘real’ whereas what is mediated will always remain essentially ‘unreal’. This perspective has been termed “the rationalist view” or “media presence” as opposed to the “psychological view” or “inner presence” (Riva 2006:73; 2008b).

Social Presence as an Experience of Others

Some scholars depict social presence as the plain experience of observing another actor, which means that social presence comes close to being the perception, awareness, sense, or feeling of the other’s mere appearance (Biocca & Nowak 2001; Mennecke et al. 2010; 2011; Swinth and Blascovich 2002). For example, Biocca and Harms (2002:11) propose that, at its lowest level, “social presence is characterized by perceptual awareness, a peripheral sense of spatial co-presence of the other’s mediated body.” Similarly, Champion (2011:74) argues that (what he calls the “etic” degree of) social presence refers to the “degree to which [someone] see[s] social interaction (mutually perceived and understood) between two or more intelligent beings.” The “passive” form of social presence, he argues, does not require active participation in social interaction—the subject can experience social presence merely by passively observing it. Mennecke et al. (2010) see social presence as the attention to other beings in the environment, but claim that the phenomenon also includes attention to the self and to the shared task. To them, social presence is the “loop of focus” in which attention is constantly oscillating between other-recognition, self-recognition, and recognition of the collaborative task. In addition, these researchers emphasize the importance of the actor’s awareness of the digitally mediated environment, and of his embodied representation in this environment.

Biocca and Harms (2002:14) suggest that other-awareness is an “automatic psychological response either small or great in which we react, model, or respond to representations of others as if they were present.” Again, the ‘as if’ in this quote reflects the idea that a certain amount of illusion is involved—to evoke a social presence response, the representations of others must trigger the actor to approach those representations in the same way they approach people that are physically co-present. Furthermore, Riva (2006; 2008a; 2008b; 2009)—who describes social presence as a phenomenon consisting of two types of internal responses to other-observation: social-presence-as-feeling and social-presence-as-process—argues that the latter response is “the continuous activity of the brain in identifying Other’s [sic] intentions within the perceptual field [...] a sophisticated form of monitoring of others’ actions transparent to the Self but critical for its social abilities”
In an earlier paper, Riva (2006:66) explains that “[t]his process is based on covert imitation: an automatic action emulator, tracking the behavior of other subjects in real-time to generate perceptual predictions.” Similarly, Lee (2004:499) argues that other-observation triggers “the mental simulation of other intelligences”; a process which is referred to as social presence.

Social presence as the experience of the other’s appearance is a definitional aspect that often co-occurs with the aspect of experienced access to the other’s internal states (see e.g. Biocca & Harms 2002, Biocca, 1997, Biocca et al. 2001; Riva 2006; 2008a; 2008b; 2009; Kehrwald 2010a; Nowak 2001). Biocca (1997:13) argues that “[o]bservers of [a] physical or mediated body read the emotional states, intentions, and personality traits by an empathic simulation of them” and that social presence is “the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another” (1997:20). In this perspective, social presence emerges as soon as the subject perceives the presence of another intelligence in a way that evokes a feeling of access to the very same intelligence (Biocca 1997; Biocca, Harms & Gregg 2001; see also Lee 2004). Similarly, Riva (2006:66) assumes that a central feature of social presence is the neuropsychological process that tracks “the behavior of the other to understand the characteristics (content and motive) of his/her intentions.” By distinguishing between “social-presence-as-process” and “social-presence-as-feeling”, Riva singles out social-presence-as-feeling as the non-mediated perception of the other’s intentions (Riva 2006:66, also in Riva 2008a). Here, the other becomes (socially) present to the subject when the subject is able to “recognize [the other] as [an] enacting” being (Riva 2008a:98), that is, when the subject experiences that s/he understands the intentions of the other. Riva goes on to claim that his view of social presence (as feeling) resembles Gallese’s (2004) “intentional attuning,” a concept that captures the ability of pre-reflexive identification with others (Riva 2008a). The feeling of empathy and communicative synchrony that characterizes social presence (as feeling) is also pre-reflexive, according to Riva (2006:66), and becomes a matter of conscious reflection only when “the quality of [the actor’s] experience is modified during a social interaction” (Riva 2008a).

Social Presence as an Experience of Reciprocity

A final way of conceptualizing social presence as an experience is in terms of a two-way connection between oneself and the other; i.e. as a subjective experience of an intersubjective connection. This perspective recurrently appears in the “short-hand” definition of social presence that depicts the phenomenon as a subjective experience of being together (see e.g. Ijsselsteijn, van Baren & van Lanen 2003; Ijsselsteijn & Riva 2003; for variations on the theme see e.g. Williams 1978; Heeter 1992; Klimmt & Vorderer 2003; Schroeder 2002; Zhao 2003; Kehrwald 2010a) or as an experience of
some form of two-way connection (see e.g. Nowak 2001; Biocca & Harms 2002; Soeffner & Nam 2007). Although in this latter perspective, social presence is most commonly discussed in terms of reciprocated awareness, some scholars also emphasize the emotional character of the interpersonal connection that the subject experiences, where subjective feelings of intimacy, togetherness, contact or participation are seen as characteristic of social presence (e.g. Soeffner & Nam 2007; Shin 2002; Champion 2011).²²

When it comes to the experience of reciprocated awareness, Biocca and Harms (2002:18) describe co-presence as a subjective sense of others and “a [subjective] sense of the others of ‘me,’” while Zhao and Elesh (2008:570) describe social presence as a state in which actors are reciprocally oriented towards each other. Similarly, Biocca, Harms and Gregg (2001:2) see co-presence as a three-staged experience that includes 1) a subjective belief of not being alone and secluded; 2) a peripheral or focal awareness of the other; and 3) a subjective experience of having others being aware of oneself. To Biocca and colleagues (2002:11; 2001:2), this experience of mutuality is fairly complex: at its highest levels, they argue, social presence involves a strong sense of intersubjectivity, such as the “subjective or intersubjective judgment of mutual accessibility of the other, such as mutual attention, mutual comprehension, shared emotional states, and interdependent behavior.” Another suggestion is that social presence constitutes “a psychological state in which the individual perceives himself or herself as existing within an interpersonal environment” (Blascovich 2002).

In these definitions, awareness is predominantly discussed in terms of other-awareness, while self-awareness goes unnoticed. However, exceptions to this rule are presented by Herrera, Jordan and Vera (2006) and Rogers and Lea (2005). Social presence, write Herrera and colleagues (2006:546), is a way of experiencing oneself through the experience of mutuality. More specifically, it is the “moment-to-moment awareness of ourselves in the process of perception of the (this time) social world, which again would include both directions of the perception/action cycle: awareness and inferences about the subjectivity of others and feeling that others are aware of us.” Thereby, the subject becomes self-aware because she is aware of the outside (social) world and feels that the outside world is aware of her.

The actor’s experience of others and/or of reciprocity is also thought to be influenced by how well she can adopt the perspective of the other, for example through empathy (Klimmt & Vorderer 2003), or theory of mind (Lee

²² There is, however, a subtle difference between the three examples mentioned. While Shin (2002) and Soeffner & Nam (2007) indicate that social presence is an emotional experience associated with a more or less significant other, Champion suggests that social presence is linked to the emotional experience of being part of social interaction, not an emotional experience of the other participants as such. Thus, the first two authors see social presence as a feeling directly associated with specific interpersonal bonds, while the latter links social presence to the emotional experience of being part of social interaction.
2004, see also Biocca, Harms & Gregg 2001; Biocca & Harms 2002; Riva 2008a; Riva 2008b). However, it is also pointed out that the actor’s intentions (Riva 2008b); expectations (Heeter 2003; Swinth & Blascovich 2002) and level of involvement and engagement (Klimmt and Vorderer 2003; Hwang & Park 2007) in the meeting with the other will make a difference to the experience of social presence.

The Influence of Relational or Interactional Factors on the Experience of Social Presence

As we have seen, social presence is rarely defined in terms of an interpersonal phenomenon (that is, an interactional phenomenon or a relation), but the literature sometimes points to links between the two, particularly when it comes to factors that may modify social presence as an experience of external events. Of course, one can argue that most of the definitions reviewed above already assume an interpersonal factor, as they often concern the availability of cues that make actors perceptible to one another, or the perception of those cues. In that sense, the majority of social presence definitions concern the establishment of an interpersonal relation between actors, either because social presence is described as the affordance that makes the relation possible, or in the sense that social presence is the resulting experience of such an established relation. However, there are also interpersonal factors that go beyond the initial establishment of social contact, and in this section I will review how those are understood in social presence theory.

In the literature, interpersonal factors are more commonly described as causes or conditions (or effects, as explained below) of the social presence experience, than they are considered parts of the phenomenon itself. For instance, it is often held that the responsiveness to social stimuli plays an important part in the emergence of social presence. Together, social stimuli and social responsiveness lay the foundation for the interplay of responses (i.e. for social interaction). As was discussed in relation to the environmental affordances, the idea that the availability of social cues in the environment will impact social presence is commonly held, and in many cases, responsiveness is acknowledged as an equally important companion to social stimuli. So not only must there be social cues available to observe, but, in addition, the observer must be aware of and responsive to those cues (Biocca & Harms 2002:17; see also Heeter 2003; and Mennecke et al. 2011). The mere existence of interaction is also identified as a factor contributing to social presence (e.g. Tu 2000, Heeter 1992, Biocca & Harms 2002:5), and previous studies claim to have shown that even the simplest form of interaction between users and technologies influences social presence (Lee 2004:495).

Several scholars suggest that the relational factors influencing social presence are not all about the interchange of stimuli and responses, but also about, for example, the nature and length of relationships between actors (Swinth & Blascovich 2002), the intimacy between actors (Rettie, 2005),
how actively actors participate (Kumar and Bensabat 2002; Mennecke et al. 2010:6p; Kehrwald 2010a), the number of participants (Schroeder 2006), and what goals they share (Swinth & Blascovich 2002; Mennecke et al. 2010; 2011). Moreover, social presence is thought to be influenced by culture (Shin 2002:126) and the meaning actors ascribe to the encounter (Swinth & Blascovich 2002:14).

Social presence theory thus acknowledges that there are social factors beyond the availability, perception and basic exchange of social stimuli that may influence the phenomenon. However, there have been few attempts at further explaining the relationship between social factors and social presence, for example, in what way culture or the shared goals of actors affect the experience of social presence.

Summary and Discussion of Social Presence Theory

Above I have reviewed the various definitions, causes, and effects of social presence described in the field of social presence research. To begin with, I chose to categorize the definitions of social presence according to what scholars assign it to. In the first category, therefore, I placed statements in which social presence is assigned to the environment; in the second category I reviewed accounts describing social presence as related to self-presentation; and in the third category I discussed statements that assign social presence to a subjective experience. The first and the third categories are closely related, while the second must be considered something of an outlier. Therefore, I will begin by a summary and discussion of the first and the third categories and return to the second one later on. Along the way, I will comment on how social presence in online chat would be portrayed in accordance with the different perspectives.

The environmental definitions of social presence either describe the phenomenon as the general affordance of social stimuli in an environment or as the specific affordance of social cues that reveal that actors are ‘there’ in the environment. In the general sense, social presence becomes a measure of how well a medium can transmit social information: the amount and variation in the social cues, but also the speed with which those cues are transmitted, are considered. For online chat, therefore, social presence would be an indication of the informational richness and immediacy that it affords, where one can assume that the disembodied, near-immediate nature of text-to-text conversation would grant online chat a rather low ‘social presence score’.

In the more specified sense, social presence denotes that the medium is able to afford ‘observability’, that is, it affords cues that signal that others and/or the actor are present in the environment. Applied to online chat this would mean that the way in which the application allows actors to appear in the environment constitutes the social presence capacity of online chat. As the visual representation of actors in online chat is rather limited compared
to, for instance, virtual worlds or video chat (where actors are typically only represented by their user name and sometimes also an optional profile picture and some additional profile information), this would again imply that the social presence capacity of online chat is lower than that of other digital environments (and of course of face-to-face environments).

The third category deals with definitions that depict social presence as an experience of, or reaction to, external events. Based on what external event social presence is thought to be an experience of, three main lines of thought can be identified. Social presence is seen as an experience of the medium, the other, or the reciprocal relationship between actors.

Overall, the described experiences vary from being a psychological reaction that is not necessarily visible to—or shared with—others, to experiences that relate to the actor’s more noticeable orientation to the outside world, such as her perception, attention and awareness of it. The definitions also diverge in a number of other ways. For one, scholars disagree about the nature of the response. Although it is most commonly described as a spontaneous and involuntary reaction, some scholars see it as a conscious and, sometimes, deliberate psychological activity. Moreover, some definitions make the concept relevant only in relation to online situations, for example when social presence is seen as the illusion of non-mediation, while others would be applicable both to online and offline situations, for instance when social presence is depicted as the psychological process through which an actor forms an understanding of the internal life of the other. This latter understanding of social presence would imply that—in theory—actors could experience the same sort of social presence regardless of whether they are interacting face-to-face or through an online chat service. The former understanding—social presence as an illusion of non-mediation—would by contrast imply that the experience is, firstly, only applicable when comparing different mediated environments to each other, and, secondly, that online chat applications would potentially be considerably less capable of providing the experience than more ‘immersive’ media (e.g. virtual worlds or video call services) are. The design of online chat applications suggests that they are not specifically created to give the user the impression that she is not using technology to communicate. While a user could probably ‘lose’ herself in a text-to-text conversation and not consciously reflect on the technology supporting it, it seems unlikely that she would ever experience that she is not engaging in a mediated form of communication. According to this definition, therefore, online chat would again rate rather low on the social presence scale.

The definitions in these two categories—the environmental and the experiential—could have been presented differently. An alternative organizing principle would have been to distinguish the accounts that put most weight on the different aspects of social perception from the accounts that put most weight on how closely the digitally ‘created’ phenomenon mimics the one
‘created’ in face-to-face environments. The first type of definition, where social presence is defined as, for example, observability, perception of others or the resulting feeling of perceiving others, does not necessarily portray digitally mediated social presence as a different phenomenon from social presence in face-to-face environments. By contrast, the second perspective departs from the assumption that ‘real’ social presence is only to be found in face-to-face environments, and that ‘virtual’ social presence only occurs in digitally mediated interaction in those cases where actors fail to notice that the interaction is mediated, or when the mediated environment affords sufficiently face-to-face-like interaction. In other words, the first type of definition is, in principle, applicable to all kinds of human-to-human encounters, while the applicability of the second type is limited to encounters in mediated environments.

What these definitions of social presence have in common, however, is that—apart from some of them taking into account what type of social information can be transmitted in interaction, and others taking into account the reciprocal process of social observation—few scholars define social presence as an interactional phenomenon. That does not, however, mean that these scholars entirely dismiss the idea that social presence would have anything to do with social interaction or social bonds, as many argue that social presence is either influenced by, or influencing, these social processes. For example, many see interaction as a contributing cause of the experience of social presence (e.g. Tu 2000, Mennecke et al. 2011, Heeter 1992, Biocca & Harms 2002:5, Lee 2004, Riva 2006) while others emphasize interaction (or aspects of it) as an effect of social presence (Riva 2006), and yet others consider social presence as an element of social interaction (e.g. Mennecke 2010)\textsuperscript{23}. But few have thus far tried to explicate this relationship and re-evaluate it in a more precise manner, and Mennecke and colleagues (2011)

\textsuperscript{23} As for the ways in which social presence affects social interaction and relationships, many scholars emphasize how social presence positively influences social interaction, most notably in terms of improving or enabling co-operation (e.g. Kehrwald 2010a; 2010b; Rice 1993; Rogers & Lea 2005, Rourke et al. 2001; Biocca 1997:20; Riva 2006:66); helping people bond and develop relationships (Kehrwald 2010a:45; Garrison 2007:63; Rice 1993, Short et al. 1976; Champion 2010; Nevejan 2007; Gunawardena 1995; Rettie 2003; Kehrwald 2010a:39p); or ensuring the establishment of a common ground (see e.g. Biocca, Harms & Burgoon 2003:459; Ijsselsteijn & Riva 2003:12; Rogers & Lea 2005:156). An individual who experiences social presence is also thought to have increased positive feelings towards the in-group and motivation to collaborate (Rourke et al. 2001; Mennecke et al. 2011:426), and to experience a more dynamic sense of others in mediated environments (Kehrwald 2010a:45). Social presence is furthermore thought to allow individuals to perceive the “content and motives of others’ intentions” (Riva 2006:70 see also 2008b; 2008a; 2009; Kehrwald 2010b:36; Mennecke et al. 2011:439) and to “interpret information related to social situations” (Kehrwald 2010b:36; see also Mennecke et al. 2011:439; and Swinth & Blascovich 2002:10). Social presence is positively associated with self-perceptions (e.g. Biocca 1997; Zhao 2005; Mennecke et al. 2011), the emergence/acquirement of a digital self (Kehrwald 2010a:45; Zhao 2005:395), and with social and personal identity (Biocca 1997; Mennecke et al. 2011:425).
similarly point out that “one of the limitations of the current literature ... is that the notion of co-presence does not take into account the role of interaction in creating a high sense of co-presence” (2011:427). One may therefore ask whether it would not, for example, be more fruitful to consider the unfolding conversation in an online chat session when examining “social presence” than to focus on the appearances of others, or actors’ sense of non-mediation in the environment?

The obvious exception from the non-interactional tendency in social presence research is of course the view of social presence as a form of self-presentation, as proposed by Rettie (2005) and in the field of online education studies. Here, social presence is the activity of self-presenting or the ability of an actor to present herself, participate, and skilfully interact in a digitally mediated situation. In online chat, an actor’s social presence would thus depend, for example, on how she chooses to appear in the environment (i.e. by her user name, profile picture, etc.), how well she can handle typical online chat jargon, and how well she manages to appropriately involve herself in the interaction. However, the burden of “creating” social presence still appears to be placed on the individual rather than on the interacting couple or group, meaning that these views also come across as having a rather one-sided focus. The online educational view on social presence furthermore differs from the others in that it does not strive to present some general theory of social presence, but rather aims to create a theoretical model that can benefit educational research. It may therefore lose some of its general applicability, because what is, for example, considered “skilful interaction” and “successful self-presentation” is probably different in an educational context compared to more generalized social contexts.

A question that is largely overlooked in the social presence literature is what the opposite of social presence—i.e. social absence—would entail. Is

---

24 But absence (i.e. not social absence), has been defined in two main ways. First, absence can be seen as pertaining to attention or lack thereof. In one of the more thorough explications of absence, Waterworth and Waterworth (2006) define the phenomenon as an abstract, conceptual process of the mind, as opposed to the concrete, perceptual process characteristic of presence. Absence is a rather reflective psychological operation in which attention is directed to “internally generated scenarios”, or figments of imagination that are “detached from the current perceptual flow” (2006:82 see also Waterworth, Riva & Waterworth 2003 and Lee et al. 2002 for similar propositions), so that the individual appears “absent minded” to others. Absence is often used to denote a type of highly unwanted allocation of attention. Thus, presence is to be attentive to the “right” environments/objects, whereas absence is to be attentive to the “wrong” ones. Here, the absence/presence distinction is not a matter of inwardly or outwardly directed attention, but a question of to what external environment/event the outward attention is allocated. As the literature reviewed here is exclusively focusing on digital environments, absence as displaced attention is most often described as a state in which individuals are perceptive of the “primary” (i.e. the physical) and not the “secondary” (i.e. the technologically mediated) environment (see e.g. Chertoff et al. 2008:409; Rettie 2004, Walker & Davide 2003; Slater 2009)
it possible to be logged into an online chat environment and still remain socially absent? If yes, is it then also possible to be socially absent while participating in an online chat conversation? Or does social absence only occur when the actor logs out of the online chat environment and puts her device to sleep? If the latter is the case, then would that not mean that merely logging in to an online chat environment equals social presence?

When the notion of social absence is at all mentioned in the social presence literature, it is mostly discussed in very vague terms. Social absence has, for instance, been referred to as an “extreme lack of social presence” (Kehrwald 2010b:37), which, due to its circularity, is a rather meaningless statement (i.e. “social presence is the opposite of social absence, which is the opposite of social presence”). Other scholars define absence not as the inversion of social presence, but as something that can co-exist alongside it. When this latter form of “social presence-in-absence” is discussed, it is usually the idea of absence as a lack of bodily cues that scholars speak of (see e.g. Leggett 2004; Soeffner & Nam 2007; Gerhard, Moore & Hobbs 2004). Scholars are here referring to social presence when speaking of presence, but bodily absence when speaking of absence (see e.g. Rogers & Lea 2005; Schroeder 2006) which means that (bodily) absence does not necessarily circumvent (social) presence but can even reinforce it, as one should not, according to Lowenthal (2010), assume that “the absence of visual cues lead to an absence of sociability” (2010:129; see also e.g. Kehrwald 2010a; Swinth & Blascovich 2002; Biocca, Harms & Burgoon 2003). Of course, it is this kind of “social presence despite bodily absence”-dynamic that originally attracted scholarly interest, but it does not shed any further light on what social absence would stand for in social presence theory. Another suggestion, that implicitly seems to point to social absence, is to describe it as a form of distant passivity, where the individual is a passive observer in complete lack of agency (Herrera, Jordan & Vera 2006; Kehrwald 2010a).

A number of scholars operating within the neighbouring field of telepresence research have noticed a similar ‘absence’ of examinations of tele-absence, and argued that this is problematic. For instance, Waterworth and Waterworth (2006) contend that “any useful definition must exclude things, and a useful definition of presence must have implications for what is not presence” (2006:83). That is because “[b]y introducing the concept of ‘absence’, presence can be distinguished from other concepts with which it is sometimes confused, including engagement, attention, or even consciousness itself” (2006:83).

Second, absence is also framed as a lack of information, specifically of information conveyed by spatial cues (Horvath & Lombard 2010), cues providing haptic feedback (Harvey & Sanchez-Vives 2005), or embodied cues (e.g. Kreijns et al. 2004).

25 Similarly, Floridi (2005) calls for an analysis of telepresence that is “able to discriminate between, and possibly explain, cases of unachieved telepresence, of failure or interruption of telepresence, of faulty or insufficient telepresence” (2005:660 see also Villani & Riva 2008).
In the social presence literature, then, being there together (which is commonly used as the short-hand definition of social presence) either denotes an environmental affordance, a subjective response or a reaction to elements in the external world, or as self-presentation. It is often related to interactional or relational elements, but it does not, in general, signify an interactional or relational phenomenon in itself. Moreover, the social presence literature does not to any greater extent examine the question of what social absence in digitally mediated environments would entail. Throughout the summary I have discussed how social presence in online chat would be understood in the light of the various perspectives, noticing that in several of them, online chat environments would be considered somewhat inadequately equipped to afford social presence.

Given that social presence is rarely defined as an interactional phenomenon I will, in the next part of this chapter, proceed to discuss what being there together could possibly mean in a more interactional context. Many (although not all) social presence scholars also hold on to the assumption that mediated social presence is only realized when it closely mimics the more ‘real’ and supposedly unproblematic way that people are being together in physical environments. Therefore, it makes sense to examine what microsociologists that focus on face-to-face meetings themselves have said on the topic of ‘being together,’ and whether they really find non-digital togetherness as unproblematic as the social presence researchers seem to do.

“Being Together” in Phenomenology and Interactionist Theory

The question of what it means to be together somewhere is not new to sociology, and has for example been discussed from the perspective of phenomenology (for instance by Schütz) and from interactionist standpoints influenced by the American pragmatism of, for example, George H. Mead and William James (e.g. by Goffman). However, and as evidenced by the emergence of social presence research, aspects of togetherness that are now posing obstacles to our understanding of the phenomenon appeared unproblematic at the time Schütz, Goffman and other classical sociologists were operating. In sociology and phenomenology, for example, the concept of presence used to have a very straightforward definition: To be present in an environment meant being physically located in a physical place; and co-presence always involved two people being physically located in the same physical space.

But even though the studies that predate the advent of digital communication technology for the most part assume that actors are physically co-located when they are together, this does not mean that ‘togetherness’ as
such is understood as nothing more than physical co-location. In fact, the literature describes how individuals in a shared physical situation can remain remote and distanced to each other, or they can be together in the sense that they do meet and overtly interact. And even within such an ‘interactive’ meeting, the literature shows how actors can be more or less distant or close.

Acknowledging that most theorizing that took place before digital communication technology entered into people’s lives was centred on face-to-face interactions does not necessarily mean that those analyses are no longer useful for our understanding of modern day social interactions. Goffman’s theoretical constructs—particularly his ideas of back and front regions and the presentation of self—have, for example, been frequently adopted in studies of digitally mediated interaction (e.g. Robinson 2007; Bullingham & Vasconcelos 2013; Zhao 2005). In an effort to adapt Goffman’s idea of framed interaction to mobile interactions, Rettie (2009) concludes that it is mostly useful for digital interactions that are relatively synchronous, while less significant for interactions that are asynchronous, either due to technological capacity, or due to the norms and expectations actors have regarding the communication. Similarly, Beneito-Montagut (2015) also finds Goffman’s notion of encounters (i.e. framed and focused interaction) to be applicable in online settings.

Goffman’s understanding of co-presence is sometimes acknowledged and credited in social presence theory as the face-to-face version of digitally mediated co- or social presence (Nowak 2001; Zhao 2003; Biocca & Harms 2002). However, the focus here remains on the availability and awareness of social cues in the environment, omitting from the discussions Goffman’s further observations on how co-presence is situated, framed and negotiated within situations. Short and colleagues’ (1976) theory of social presence has also been likened to Goffman’s idea of the presentation of self (Tu 2000; Rettie 2005), as both theories emphasize the importance of information that is expressed (as opposed to communicated) in interaction. However, Rettie (2005) also points out that Short and colleagues do not, like Goffman, consider how the socialized actor himself does self-presentation, but rather depict social presence as a straightforward effect of what cues the medium affords. While also discussing social presence in relation to self-presentation, the online educational perspective reviewed in the former part of this chapter does not draw on microsociological ideas and concepts in their theorizing.

In the following I will firstly explore what the ‘basics’ of togetherness could be said to stand for in interactionist and phenomenological literature. What are, for example, the minimal requirements for actors to ‘be together,’ and what shapes could ‘minimal’ togetherness take? Thereafter, I go on to examine what the literature has to say about degrees of togetherness, more specifically about what makes actors closer or more distant from each other in togetherness.
From Independency to Interdependency

One way of capturing what ‘being together’ would mean in the literature is to look at how it describes that very moment when independent actors enter into a relationship of interdependency with others. Of interest is therefore the minimum number of elements needed for such a relationship to emerge.

Being in the social world, Schütz (1932/1967:146) writes, assumes a certain social behaviour of the actor, namely other-orientation. It is an attitude directed towards the other person’s “duration”, where the other is seen as a conscious and experiencing subject. The nature of other-orientation can, however, vary greatly. For example, it does not require reciprocation on the part of the other and may remain a purely internal happening: it can exist and continue without the other taking on the same attitude, and it does not require any communicative or expressive action. However, when Schütz (1932/1967) goes on to describe the “orientation-relationship” he also claims that, in its non-reciprocated form, it is disqualified as a true social relationship.

Schütz’ definition of a “living social relationship,” by contrast, involves one actor immediately reciprocating the ‘other-orientation’ of another actor, while both are fully aware of this reciprocity of orientation. Thereby, actors are “living in a mutually related conscious experience” and “carried from one moment to the next in a particular attentional modification of the state of being mutually oriented to each other” (1932/1967:156).

In a similar vein, Goffman (1963:17) defines co-presence as a state where individuals are physically co-located and experience each other with their naked senses (meaning that information sent and received is embodied and immediate). This means that actors “sense that they are close enough to be perceived in whatever they are doing, including their experiencing of others, and close enough to be perceived in this sensing of being perceived.” Since co-presence “renders people uniquely accessible, available, and subject to one another” (1963:22) it follows that individuals who share a physical environment are not automatically co-present. For instance, a person located on a public street might not be able to reciprocally observe each and every other person on the street, and is therefore not in a relationship of co-presence with those people. Instead, co-present individuals must share a situation, which, as described in Chapter 2, emerges when actors are mutually accessible to one another. Such ‘gatherings’ can still represent a very loose social form, in

---

26 The “orientation-relationship” is established as soon as an actor takes on the attitude of other-orientation. In other words, an orientation-relationship exists when a person apprehends another person as an experiencing social being with an inner life. The “orientation-relationship”, Schütz (1932/1967:155) writes, is difficult to objectively observe, as the social behaviour it is constituted of need not be overt (e.g. it can consist of social feelings directed towards the other).
that even if a new person enters the situation it does not necessarily cause any apparent change in what the first person is doing.

Goffman (1963:24) goes on to suggest that the behaviour of co-present actors takes two different overarching forms: either unfocused or focused. Unfocused interaction represents a form of subtle and almost peripheral monitoring where actors are drawing information about each other by, for example, throwing brief glances. Therefore, Goffman argues, the only purpose this behaviour serves is to manage co-presence. By contrast, focused interaction—also called an encounter or face engagement—occurs when actors, for example through conversation, “openly cooperate to sustain a single focus of attention” (Goffman 1963:24). Focused interaction thereby represents a mutual activity from which other co-present individuals can be excluded (1963:83).

So how do actors move past the initial recognition of others and the monitoring of their co-presence, and into a mutual activity—that is, how does unfocused interaction transform into a focused encounter? As Goffman (1963:91p) describes it, a transition from unfocused to focused interaction is typically initiated by a subtle agreement between actors, where (because of its immediacy) eye contact holds an especially important place.

An encounter is initiated by someone making an opening move, typically by means of a special expression of the eyes but sometimes by a statement or a special tone of voice at the beginning of a statement. The engagement proper begins when this overture is acknowledged by the other, who signals back with his eyes, voice, or stance that he has placed himself at the disposal of the other for purposes of a mutual eye-to-eye activity—even if only to postpone his request for an audience.

In their discussion of the opening of a social act, Miller and colleagues (1975) similarly argue that the very first thing which needs to happen is for actors to acknowledge that the attention they direct to each other is reciprocal. This acknowledgement can be tactile or visual (in which case it is often simultaneously reciprocal), or oral or written (in which case reciprocity is managed sequentially. For example, Actor 1 says “Hello”, whereby Actor 2 reciprocates with a “Hi”). If in observing the other, the other is not reciprocating this observation, then this would not suffice to even begin opening up for social interplay.

27 Instead of describing the activity characterizing these phases in terms of social acts, Miller and colleagues (1975:480) insist we use the term ‘concerted behaviour’ to describe what is going on. That is because the term ‘social act’ “implies a discrete entity temporally bounded by markers” which is not always the case in the three phases of an encounter. Rather, the behaviour of an opening, a middle, or a closing is continually being constructed, maintained, disintegrated and reconstructed. Although this suggestion is reasonable, I have not adhered to it in the present work, as, for my purposes, I do not consider that level of specification to be necessary.
According to Miller and colleagues, reciprocal other-orientation is necessary to, but does not suffice for, the opening of a social act. In an opening, Miller and colleagues argue, the acknowledgement of mutual attention (i.e. the reciprocal other-orientation) must firstly be followed by mutual responsiveness. That is to say that the actor builds her acts “off the prior, simultaneous, or anticipated act of another, and in the process of so doing informs the other participant(s) of that fact.” The actor is thus, in part, organizing her behaviour according to the other party’s behaviour. In addition to establishing mutual responsiveness, actors must also develop congruent functional identities, meaning they establish a relation and project (if only for a brief moment) a shared future. In those cases where actors are doing something with each other (i.e. when an object or event is attended to by two or more parties with each aware they are attending to the same object or event) a further element also needs to be added to the opening—namely a shared focus (Miller et al. 1975:483). The authors argue that only after these elements—reciprocal acknowledgement, mutual responsiveness, congruent functional goals, and (in some cases) shared focus—are established can the opening of a social act proceed into its middle, and, later, closing phases.

Miller and colleagues’ addition of congruent functional identities and shared focus points to a topic that has been discussed at length in the past. This topic concerns actors’ shared understanding of what is going on in the meeting, which among other things would include congruent functional identities and shared focus. An actor’s understanding of what is going on (often referred to as her definition of the situation) will shape what she expects of herself and others, and of course also how she interprets and draws meaning from the happenings. While there is no way an actor can literally reach into the mind of the other and confirm that her understanding is truly shared, actors can infer that their understanding of the meeting is mutual from the information they draw from the behaviour of others. Even in very unfocused situations, an actor will have some idea of what is going on and what is expected of her. She will, for example, attempt to manage her appearance and her bodily expressions when she is walking down the street, even if she cannot confirm that she is in fact being observed or noticed by other actors. Similarly, Schütz (1967:166p) states that even when there is only a basic reciprocal awareness between actors, the actors will still form hypothetical understandings of each other. In turn, these understandings are based on an understanding of the shared immediate situation. Even the most elementary we-relationship thus denotes a state in which two people share a ‘meaning context’ that helps actors understand each other as well as what is going on in the meeting.

Goffman (1974:21) uses the concept ‘primary framework’ to refer to the actors’ understanding of what is going on in a situation. As soon as the actor recognizes an event, she will interpret the meaning of this event according to the primary framework she applies to the situation. The organization of a
framework can vary in degree: some are highly structured while others are more inarticulate, merely providing a rather diffuse perspective on what is going on. In either case, the framework is both a tool for interpreting and applying meaning to whatever occurs in the situation, and a tool for deciding what is the appropriate and expected way of behaving in it. A person is also likely to employ several different frameworks to the situation. For example, her belonging to a certain group may lead her to apply the primary framework of this group (a form of cultural framework) to the situation; at the same time, she may also use a more loosely organized framework to make a specific interpretation of what is happening in the situation. While the framework may lead actors to draw faulty conclusions, Goffman also argues that actors are highly skilled at employing frameworks and ordinarily do not apply ones that are completely misplaced. However, if an actor realizes that he has misframed the situation he will likely try to break away from this inefficient frame and employ a more supportive one. Breaking frame (as well as misframing) is, however, something that typically leads to negative experiences among participants and may, when it occurs, even cause the situation to collapse (Goffman 1974:357). An important point to make here is that Goffman does not ‘just’ see actors’ frameworks as rather abstract informational contexts, but portrays them as tools that are actively employed in the situation and that are also to some extent co-constructed and negotiated in the situation (i.e. actors can also be said to be ‘doing’ framework). Frameworks thus play an active role in shaping social experience.

So to Goffman, the understanding of what is going on is determined by the frameworks employed, and these frameworks help the actor understand what is going on, both with regard to the more stable and general social structures, and with regard to the unique structures emerging within the specific meeting. One implication of this is that in both unfocused and focused interaction, the actor’s understanding of the situation will shape how she makes herself accessible (or co-present) to others, as different frames come with different rules for what is appropriate accessibility. This indicates that actors only need to employ rather generalized frameworks to understand what is going on in unfocused interaction (where actors are doing little more than managing co-presence), while focused interaction might require more specific frameworks.

The cited literature suggests that the minimal requirements for actors to ‘be together’ are that actors are mutually aware of—or co-present with—each other. But in order to be together in a more focused way, actors must sustain mutual responsiveness and attention to each other, and also share some elementary understanding of what is going on. So how would the literature describe focused interaction—or encounters—that ask for little more than that these basic requirements are met?

Sociability (or ‘play’ as it is also called in the literature) is one example of an elementary encounter described in the literature. Georg Simmel
Simmel (1910/1971) would argue that sociability is an autotelic activity that actors impulsively engage in. As such, it has “no ulterior end, no content, no result outside of itself” and it is not motivated by external, instrumental goals. Sociable activity is thus something actors do for the sheer sake of doing it and they obtain satisfaction from being associated with others and from having their own solitariness “resolved into togetherness, a union with others” (Simmel 1910/1971:128). This resembles Asplund’s (1987) idea of social responsivity (“social responsivitet”), which is described as a fundamental and elementary form of human sociality. Asplund argues that humans are wired to respond socially to stimuli, and that, in its elementary form, this innate responsiveness is not predetermined, but instead improvised and spontaneous. By using the example of a child playing with a kite, Asplund (1987:36pp) goes on to illustrate how the social responsivity of humans is so strong that almost any stimulus can entice a response. As the wind grabs the kite, the child will experience the string pulling his hand as a social invitation and respond by for example letting out a little more string. The whole relationship between the kite and the child revolves around this chain of simple stimuli and responses (something Asplund, 1987:52, calls a “responsorium”). At the other end of the continuum we find asocial non-responsivity (“asocial responslöshet”), which signifies a complete lack of response to stimuli (1987:12). There are several reasons an actor would not respond to stimuli: the lack of response can either have non-voluntary origins (e.g. the actor may have grown tired and uninterested in the stimuli, or may have been distracted by other stimuli and thus directed her responsiveness away from the original stimuli), or it can be something she has learned through socialization. A significant part of socialization revolves around taming the child’s social responsivity by teaching her when it is appropriate to respond to stimuli and when it is not—she may for example be instructed to ignore strangers but remain responsive to her friends and family. The day-to-day social interactions of adults thereby represent a more disciplined social responsiveness that has lost some of its original spontaneous and undiscriminating character.

To return to Simmel’s (1910/1971:136) notion of sociability, he considers free interaction to be one condition for sociability, which means that, overall, the social activity must be improvised and largely unregulated, where what happens next in in the situation cannot be controlled or easily predicted. “In sociability”, he argues, “talking is an end in itself; in purely sociable conversation the content is merely the indispensable carrier of the stimulation, which the lively exchange of talk as such unfolds”. In other words, sociable communication can be repetitive nonsense carrying little meaning other than the meaning necessary to sustain sociability, but it can also appear to be highly sophisticated conversation, that, for an onlooker, would seem to carry plenty of “additional” meaning. It would appear therefore, that—to Simmel—what type of social cues an environment can afford is less important than
that it can afford social cues. Instead, the exchange of virtually any stimuli and responses could be meaningful enough for sociability, and will not be affected by the lack of certain types of cues or the abundance of others. However, another implication of Simmel’s description of sociability is that, no matter whether social stimuli are available or not, the activity will be repressed when interaction becomes too regulated and scripted.

“Free interaction,” as Simmel chooses to phrase it, should therefore not be interpreted as interaction completely deprived of any form of structure, or any type of link between actors. In fact, for sociability to occur there must exist some kind of shared understanding of what is going on in interaction, and in this case the understanding of interaction as ‘free’ or ‘sociable’ constitutes the structure. In describing sociability as a type of framework, Goffman (1961:21) emphasizes precisely this, in that he is suggesting that even the rather unruly social activity characteristic of ‘sociability’ is organized. Moreover, the notion of sociability as a frame also helps explain why the activity seems to occur more frequently in certain types of social settings, while, in principle, it could equally as often occur in all types, regardless of whether they are ‘serious’ (e.g. a business meeting) or ‘unserious’ (e.g. a dinner party).

Situations in which sociability are more likely to occur Goffman (1963:132) calls open regions, where “any two people, unacquainted or not, have a right to initiate face engagement with each other for the purpose of extending salutations.” As Oldenburg and Brissett (1982:269) point out, some public settings—so-called third places—are more institutionalized open regions. Sociability is, for instance, more frequently encountered in the pub or the park than it is in the work place. That is because third places—like the pub or the park—put very loose or unspecific demands on how interaction should be organized, and actors typically “gather there primarily to enjoy each other’s company.” Parenthetically, due to the comparatively loose organization of social behaviour in digitally mediated social settings, they have often been likened to third places (e.g. Steinkuehler & Williams 2006; Moore et al. 2009; Rao 2008; Duchenaut et al. 2007)28.

However, others have pointed out that sociability can survive on less structure than a shared frame. For instance, Cooley (1902:87) proposes that

---

28 However, it has also been pointed out that digital third spaces often differ from Oldenburg and Brissett’s (1982) definition of the term. For example, Soukoup (2006) points out that many online environments—like discussion forums—are exclusive and narrow in focus, which goes against the inclusive and general interaction of Oldenburg and Brissett’s third places. In addition, although online places may be insensitive to ‘traditional’ status markers, like age or occupation, they are often hierarchically organized, where inexperienced ‘newbies’ gain the lowest status. Steinkuehler and Williams (2006) could also observe that regular users of massively multiplayer online games (MMOs) tended to engage in ‘third place-like’ behaviour only in the initial phases of their activity in the environment. As users gradually became more familiar with, and involved in, the social environment, they would take on responsibilities and obligations that made their activity more similar to work than play.
the experiencing subject does not have to have insights of the emotional states of others, or sympathy (which, in Cooley’s terminology, translates into a deeper understanding of the internal states of the other) in order to be sociable. He writes:

A child who is extremely sociable, bubbling over with joy in companionship, may yet show a total incomprehension of pain and a scant regard for disapproval and punishment that does not take the form of a cessation of intercourse. In other words, there is a sociability that asks little from others except bodily presence and an occasional sign of attention, and often learns to supply even these by imagination.

What Cooley describes resembles what occurs in the first two phases of an opening (as explained above): for sociability to occur, the other must be revealing social cues, be observed, and potentially also return the observation. This idea also goes in line with the social responsorium that can emerge between a boy and his kite, as described by Asplund. However, to Cooley, this observation of the other, and the responses of the other, can also be entirely imagined.

In a related vein, Jerolmack (2009) uses the example of play between humans and animals to argue that the importance of intersubjectivity in elementary social forms is over-estimated. Even though it may improve the interaction, neither actual nor assumed mutuality of understandings or intentions are necessary for coordinated interaction. This means that the understanding of an encounter as playful may be unevenly distributed between actors, for example in the interaction between a dog and his owner, where the owner frames the interaction as ‘play’ while the dog might just see it as an opportunity to get treats. Jerolmack (2009) goes on to suggest that what is really required of the participants is a mutual orientation to the life world and that their respective goals are compatible for play.

Similarly, Goffman (1963:20) argues that participants in a social occasion (like a cocktail party) may frame the occasion very differently yet still share an overarching understanding of what is going on. For instance, the role that the waitress takes on requires her to frame the occasion as work, while the guest may frame it as play or at least as an informal event. Yet, both the waitress and the guest will also frame the situation as a cocktail party. Thus, multiple social frames can be at work in one and the same setting without disturbing the overall order. However, unlike Cooley and Jerolmack, Goffman still appears to assume that, on some level, actors still have an understanding of the situation in common, even if the more specific frames imposed by their respective roles are different.

As described in the literature, sociability is a very elementary—sometimes even labelled the elementary—social form, which requires actors to recognize others and be responsive to them. It is also an activity that is
described to require some level of organization, if only to a very minimal degree. Sociability is said to originate from an innate social impulse in humans, which can be likened to a drive to connect socially with others. Another example of a ‘basic’ type of social encounter that can be found in the literature comes across as the complete opposite to sociability, in that it rather serves to distance actors from one another. The type of encounter I am speaking of is what Goffman (1971; 1963) calls “civil inattention,” and it signifies the act of acknowledging the existence of the other (e.g. by keeping brief eye contact), and then quickly shifting attention away from him or her.

According to Goffman (1971), civil inattention is predominantly employed in public places to ensure that life runs smoothly (e.g. that people in a crowded place do not bump into each other) at the same time as actors stay anonymous and socially inaccessible. That way, people can peripherally observe (and be aware that they are observed by) others, while still keeping at a social distance. Here, actors engage in behaviour characteristic of an opening for the sole purpose of ensuring that the interaction does not proceed into a middle stage. This type of social encounter is focused (Goffman 1963:83p), because it requires of actors that they (however minimally) recognize and respond to each other, and that they share some understanding of what is going on.

So far, I have discussed different forms of ‘minimal’ togetherness, as described in the literature. While the reciprocal attention required for living social relationships and co-presence can be sustained by means of unfocused interaction, additional elements are needed for the togetherness to become ‘focused’. Taken together, mutual recognition and responsiveness could be said to tell people that they are being together, and the shared understanding, or frame, tell people how they are being together. I have also described two types of encounters that the literature discusses, and have explained that both could be said to be representative of elementary focused togetherness, namely sociability and civil inattention. However, although these types of encounters are basic, they are also very different from each other: while sociability is an activity that draws people closer together, civil inattention is an activity that ensures a certain social distance. There are surely other types of social meetings that also represent a form of basic togetherness. But the two selected examples neatly point to what will be discussed in the remainder of this chapter, namely the various ways in which the literature describes how actors get closer to, or further away, from each other while they are being together.

Closeness and Distance in Social Experience
The discussion so far has already touched upon the topic of closeness and distance among actors that are together, for example in outlining the shift from unfocused interaction to focused encounters, and in describing how
some encounters seem to be characterized by a greater closeness between actors than others. In the remainder of this chapter I will examine this topic further, beginning by considering Schütz’ (1932/1967) distinction between direct and indirect social experience.

To Schütz, social experiences can be arranged on a scale ranging from direct (i.e. spatiotemporally immediate), to indirect (i.e. spatiotemporally non-immediate) experiences. A person we experience directly Schütz calls a “fellow man” (or a Thou), while a person we only experience indirectly he calls a contemporary (or a They). However, the exact moment when a spatiotemporally immediate Thou-relationship transforms into a They-relationship, or vice versa, is not entirely clear cut:

The first steps beyond the realm of immediacy are marked by a decrease in the number of perceptions I have of the other person and a narrowing of the perspectives with which I view him. At one moment I am exchanging smiles with my friend, shaking hands with him, and bidding him farewell. At the next moment he is walking away. Then from the far distance I hear a faint good-by[e], a moment later I see a vanishing figure give a last wave, and then he is gone. It is quite impossible to fix the exact instant at which my friend left the world of my direct experience and entered the shadowy realm of those who are merely my contemporaries. As another example, imagine a face-to-face conversation, followed by a telephone call, followed by an exchange of letters, and finally messages exchanged through a third party. Here too we have a gradual progression from the world of immediately experienced social reality to the world of contemporaries. In both examples the total number of the other person’s reactions open to my observation is progressively diminished until it reaches a minimum point. (1932/1967:177)

Although Schütz maintains that all direct social experiences—and thus all Thou-relationships—occur within the face-to-face situation, his discussion of the difficulties in drawing an exact line where an experience goes from being direct to indirect suggests that his idea of face-to-face interaction involves more than just a situation where the bodies of two individuals are located in close physical proximity. As a common procedure is to depict online interaction as the opposite of face-to-face interaction, it is important to acknowledge that, to Schütz, many forms of digitally mediated interaction could potentially still fall within the category of direct Thou-relationships. For instance, he argues that a phone call still falls within the realm of direct experience, and thus the phone-callers have a Thou-relationship—yet it is less concrete than when two people meet over a cup of coffee. By contrast, Goffman (1963:17) states that the full conditions of co-presence must involve actors perceiving each other with their naked senses and perceiving that this perception of the other is being reciprocated. Therefore, an actor may experience another person down the street with his naked senses, yet not be in a relationship of co-presence with him. The determining factor for Goffman is thus the mutuality of the observation, and the perception of this
mutuality, while, for Schütz, concreteness is determined by how ‘richly’ the other can be perceived. To understand this it is necessary to flesh out what Schütz included in the concept of They-relationships.

**Indirect Social Experience**

The main difference between Thou- and They-relationships is that in the world of contemporaries, “we never encounter real, living people at all” (Schütz 1932/1967:205). In They-orientation, the actor is not apprehending a concrete individual and does not have access to the ongoing flow of his consciousness. I know that a contemporary co-exists with me in time, but I do not have direct access to him. Therefore, “his subjective experiences can only be known in the form of general types of subjective experience […] my whole knowledge of [him] is mediate and descriptive” (Schütz 1932/1967:181p). In other words, the actor apprehends the contemporary as a social type based on his pre-existing knowledge of the social world and of social beings.

One important implication of this is that the accuracy with which an actor can make inferences about what is going on inside the head of the other is compromised, because here the actor is making her assumptions of the other based on different types of second-hand experiences. So whereas the actor, in a Thou-relationship, comes to understand the other through directly and subjectively experiencing him (and thereby also has the ability to continuously check whether her understanding of him corresponds to what she directly experiences, and revise it when necessary), she must make use of everything but her first-hand, direct experience to make inferences about her contemporaries. As an effect, she can only retrospectively discover whether her understanding of the contemporary was correct. Of course, even when she directly experiences an other, she still does not have full access to what is going on inside his head, which means that even here there is a risk of misunderstandings. However in the They-relationship, this risk is significantly increased:

A social relationship between contemporaries consists in this: Each of the partners apprehend the other by means of an ideal type; each of the partners is aware of this mutual apprehension; and each expects that the other’s interpretive scheme will be congruent with his own. [---] Here, each partner has to be content with the probability that the other, to whom he is oriented by means of an anonymous type, will respond with the same kind of orientation. And so an element of doubt enters into every such relationship. (Schütz 1932/1967:202)

As this non-immediate other-experience is a defining characteristic of the They-relationship, the actor cannot turn directly to the contemporary to confirm that her understanding of the contemporary is correct. If she could, her They-relationship would already have been transformed into a Thou-
relationship. Therefore, a They-relationship forces her to rely much more heavily on assumptions about the other, and about her and the other sharing an interpretive scheme.

As for further differences between an actor’s relationships to fellow men and contemporaries, respectively, it is firstly so that in the latter relationship her knowledge of the other remains comparatively fixed, while in a Thou-relationship she is constantly broadening and revising her understanding of the other. Secondly, the actor cannot assume that the way she uses a particular sign system is identical to how the other uses it. The more anonymous the contemporary of an actor is, the more cautious she has to be about how she uses certain words and phrasings in indirect communication (1932/1967:204). For example, in a letter to the authorities she must take care to use the shared language as objectively as possible, since she cannot detect and correct misunderstandings as immediately as she can in face-to-face interaction.

As just indicated, the second-hand experiences an actor uses to make sense of her contemporary still vary in terms of how adequately they allow her to understand the other and reassure herself that interpretive schemes correspond. Of all the contemporaries in her world, some are rather concrete to her while others are completely anonymous. In short, the more she relies on objective rather than subjective meaning contexts to make sense of her contemporary, the more anonymous the contemporary is to the actor, and the greater is the probability that the actor’s understanding of the other will not correspond with reality.

To take an example, my mother currently lives within my world of contemporaries, and yet there is a considerable difference in how anonymous she is to me, compared to how anonymous my “habitual” contemporaries (such as the police officers patrolling the city I live in) are. That is because the knowledge I have of my mother is what Schütz (1932/1967:196) calls “characterological,” which means that, although my current experience of her is indirect, it is still heavily based on experiences drawn from our previous Thou-relationships. Schütz also maintains that an actor can get a characterological understanding of a given contemporary by having an intermediary (who has met this particular person as a Thou) describe the contemporary to the actor. Habitual contemporaries, by contrast, are defined exclusively in terms of their functions (as ‘police officers’ and ‘post men’ are) and the actor’s knowledge of them is thereby more generalized. Moreover, characterological contemporaries are not only closer to our direct experience in a retrospective fashion, but they also constitute a group of contemporaries that may with very little effort be converted into fellow men (1932/1967:195pp). For example, my mother may be a contemporary to me for most of the time, but I can easily and pretty much at any time arrange for us to directly meet.

In what could be said to be an analysis of the ‘most close’ people can get, Mjöberg (2011:85) points to intimacy as a form of intense shared presence
(närvaro). To her, intimacy is not one way of relating to the other and the relationship, but rather encompasses four distinct modes of relating. Interestingly enough, one of the four modes Mjöberg mentions is actually a form of relating that can occur between an actor and his contemporary.

By using the example of letter writing, Mjöberg (2011:81) illustrates how the relational mode of ‘mental closeness’ (närhet) is not entirely dependent on an immediate stream of information between actors, but that it is rather an experience that can occur even between contemporaries that are spatiotemporally distant from each other. When a person writes a letter to his old friend, he reveals himself, his intentions, and feelings to the distant other through the words and sentences he chooses to put down on paper. In so doing, he is directed towards his mental representation of the friend, and thus a sense of intimacy can occur between the actor and the representation. So even though the friend that serves as the model to the mental representation is physically (and probably also mentally) absent from the entire activity of relating, the actor can still experience a certain closeness to him. The actor is therefore both directed towards and responsive to someone who is spatiotemporally distant. When the distant other later receives the letter, she will in turn be directed towards her mental representation of the actor and in a sense share the actor’s directedness as it is communicated in the letter. Therefore, the closeness experienced during this mode of relating is not necessarily shared in the moment—and even if one actor’s mental closeness to a spatiotemporally distant other is, in fact, reciprocated in the moment, the actor will not receive any evidence of that fact until much later (if at all).

The example of letter writing between friends or lovers also suggests that Mjöberg does not believe that (mental) closeness can occur between any two contemporaries, but that it rather happens between contemporaries that, in Schütz’ terms, would have a ‘characterological’ understanding of each other, such as friends or lovers do. The extent to which an actor can ‘closely’ relate to the representation of another person would thereby depend on, for example, what experiences and memories this representation is made up of, and perhaps also the emotions and importance that are associated with the contemporary. Although the idea of closeness is exemplified by situations in which the actor is not sharing a spatiotemporal location with the person she is relating to, Mjöberg (2011:80) does indicate that mental closeness might be an important feature of intimacy in more direct social experiences as well. The flexibility of this concept points to its usefulness for the analysis of closeness and distance in digitally mediated encounters, where actors are not as distant from each other as true contemporaries are, yet not as closely together as fellow men (described below) are either.

This last sentence points to a problem with how Schütz bases his classification of social experiences on spatiotemporal immediacy (or the complete lack thereof), where in the one realm actors are directly, or spatiotemporally immediate, experiencing each other, and in the other they are experiencing
each other indirectly in spatiotemporal distance. While lacking in spatial immediacy, digitally mediated interaction can be near-synchronous, meaning it is not entirely devoid of temporal immediacy. It stands to reason, therefore, that social relating in digitally mediated environments is not quite a relationship between contemporaries, neither is it a relationship between directly experienced fellow men.

In an attempt to tackle this shortcoming in Schütz’ theory, Zhao (2004) suggests that cyberspace has opened up a third realm of experience—the realm of consociated contemporaries (which is the same as to say contemporary fellow men). While the underlying idea is good, the proposed model is problematic in that Zhao’s understanding of contemporaries and fellow men appears not to entirely correspond to what Schütz wanted to express by the terms. ‘Consociated’ here comes to stand for an immediate relationship between actors (and it may be noted that Zhao grossly exaggerates the immediacy of digitally mediated interaction29), while ‘contemporary’ stands for an anonymous other. Zhao’s concept of consociated contemporaries therefore only encompasses people who “interact online with one another under aliases or pseudonyms in disembodied text mode to conceal their true identity” (2004:101) making the discussion a matter of how actors protected by anonymity engage in personal interactions on the Internet, and leaving out all other ways in which actors relate to each other in digitally mediated environments. While a contemporary, to Schütz, is not necessarily someone who is anonymous to the actor (as in the example of characterological contemporaries, such as spatiotemporally distant family members), Zhao argues that digitally mediated interaction between family members is simply a case of interaction between fellow men. But as has been illustrated above, Schütz’ distinction between contemporaries and fellow men has nothing to do with what actors know or do not know about one another (even though one can distinguish between different types of contemporaries based on such knowledge), but about how they experience each other in the current moment (directly or indirectly).

The problems with Zhao’s model aside, he does make a valid point in arguing that the digitally mediated environments allow actors to experience each other in a way that is neither indirect nor entirely direct. This insight will be useful in later chapters of this dissertation, but for now I wish to proceed to discuss spatiotemporally immediate—or direct—social experiences.

---

29 This is mostly due to a neglect of the role spatial immediacy has on overall immediacy in the encounter, where Zhao suggests that the temporal immediacy afforded by digitally mediated environments is on a par with spatiotemporal immediacy. This results in questionable conclusions such as: “Individuals share a community of time because they are able to maintain simultaneous contact with one another, allowing the streams of their consciousness to flow side by side from moment by moment. Through instant messaging, for example, distant individuals can hold an instantaneous conversation in writing, expressing opinions and emotions, debating issues or flirting with each other. In the temporal simultaneity of the exchange of thoughts and feelings, these individuals ‘grow older together’” (2004:99).
Direct Social Experience

On the most direct end of Schütz’ (1932/1967) continuum, we find pure Thou-orientation, which occurs only in the face-to-face situation. As was mentioned earlier, Thou-orientation does not, in its purest form, involve the specific characteristic of the other or any judgments of what he or she is thinking about—it is merely an awareness of the co-existence of another social being.

Thou-orientation may be reciprocal (in which case Schütz calls it a we-relationship) or one-sided, which also means that, unlike in They-orientation, there is a difference between participation and observation in the world of fellow men. When the Thou-orientation is non-reciprocal, the actor is aware of the other without that other being aware of the actor in return. It thereby lacks the “mutual mirroring characteristic” of reciprocal Thou-orientation, where actors simultaneously assess what is going on in each other’s minds (Schütz 1932/1967:173). In either case, the directness of experience is essential. Even if I am only peripherally observing someone standing next to me in a crowd, and even if I only have a very inadequate understanding of him, my orientation still has a Thou-character.

“[I]n the face-to-face situation, directness of experience is essential, regardless of whether our apprehension of the Other is central or peripheral and regardless of how adequate our grasp of him is. I am still ‘Thou-oriented’ even to the man standing next to me in the subway.” (1932/1967:176)

What Schütz (1932/1967) is saying is thus that, even though the concepts of Thou-orientation and the we-relationship are typically used to describe more concrete relationships, the terms can also be used to describe “the lower limits of experience obtainable in the face-to-face relationship, in other words, for the most peripheral and fleeting kind of awareness of the other person.” This latter type of behaviour is reminiscent of Goffman’s earlier mentioned concept of unfocused interaction, which represents a behaviour that does little more than monitor actors’ co-presence in a shared environment.

However, as was mentioned earlier, the non-reciprocated form of Thou-orientation does not, according to Schütz, qualify as a living social relationship, which is why I will not explore this type of Thou-orientation further. Similarly, I will pay more attention to focused than unfocused encounters, as the former can be understood as a more concrete we-relationship than the latter.

In reciprocal Thou-orientation—the we-relationship—participants are “sensitively aware of the nuances of each other’s subjective experiences” (Schütz 1932/1967:202). But even here, relationships come with different degrees of concreteness, or closeness to direct experience, and it is the varia-
tion of concreteness (or closeness) within we-relationships (or focused encounters) that I will discuss in the following.

Schütz (1932/1967:168) suggests that the concreteness of the other and of social relationships depends on how actors ‘know’ each other. However, how people know each other should, in this case, not be confused with what people know about each other. Although the amount and type of knowledge one has of the other is what determines whether she is a characterological or habitual contemporary in indirect experience, it is the way of knowing that decides how close or distant actors are to each other when they are directly together in a we-relationship.

Schütz (1932/1967) identifies four different variables, each of which presumably stands for a distinctive way of knowing the other, and that—in their own respective ways—influence how ‘concrete’ actors are to each other. These are immediacy, point of view, intensity, and intimacy. As was mentioned earlier, immediacy and point of view have to do with how the environment affords the actor the ability to perceive the other, and are thus responsible for whether or not actors can at all be (directly) together. Temporal immediacy here means that the actor is able to be other-oriented and at the very same time apprehend the other’s other-orientation. There is no time delay and no question that has to await an answer—question and answer happen simultaneously. Thus, no matter what actors do, the knowledge they have of one another is growing from moment to moment when the relationship is immediate. However, even within the realm of direct experiences, immediacy can vary, for example because of variation in the point of view from which the other is experienced. In a room full of people, if a person is standing (from my point of view) at the back of the crowd I will experience that person less immediately than someone standing right in front of me. That is because point of view affects spatial immediacy, even if it does not necessarily affect temporal immediacy. When Goffman describes how co-located actors on a public street may not be co-present to each other it is not because the environment does not at all afford immediacy, but because the point of view from which the actors are observing each other reduces the spatial immediacy between them (which in turn makes mutual monitoring impossible).

While immediacy and point of view concern how the environment allows actors to ‘know’ each other, intimacy and intensity are ways of relating or approaching the other that further modify the access to knowledge of that other. However, Schütz (1932/1967:168p) does not explicate exactly what he means by ‘intimacy’ and ‘intensity,’ but instead settles for the following description:

Compare, for instance, the knowledge two people have of each other in conversation with the knowledge they have of each other in sexual intercourse. What different degrees of intimacy occur here, what different levels of con-
sciousness are involved! Not only do the partners experience the We more deeply in the one case than in the other, but each experiences himself more deeply, and his partner more deeply. It is not only the object therefore, that is experienced with greater or lesser directness; it is the relationship itself, the being turned towards the object, the relatedness.

These are only two types of relationship[s]. But now consider the different ways in which they can actually take place! The conversation, for instance, can be animated or offhand, eager or casual, serious or light, superficial or quite personal.

What Schütz is describing here is how the general way in which actors ‘know,’ or relate to, each other in different types of encounters (sexual intercourse, conversation, sociability, civil inattention, and so on), also means that actors are generally more closely together in some encounters, and more distantly together in others. But within each type of encounter, there is also room for variation in how close actors get, depending on how actors actually relate to each other (e.g. in a casual and superficial way, or in an animated and personal manner). This resembles the relationship between a frame and the framework actors ‘do’ within it, where the ‘type’ of encounter would represent the frame and the variation would represent the negotiation of the frame (i.e. as the activity of framework).

Precisely how Schütz believes that intimacy and intensity (as ways of relating to the other) work to bring people closer together remains rather obscure, yet the overall point he is making—that ways of relating affect closeness—is clear. In the following, therefore, I examine different ‘modes’ of relating described to have a bearing on the closeness or distance in encounters.

**Intensity/Involvement**

Although Schütz does not specify what he means by the intensity of an encounter, one possible interpretation is that intensity refers to a certain degree of involvement. As it is described by Schütz (1932/1967), more concrete (or close) we-relationships come across as encounters in which participants are intensely involved. Here, actors are reciprocally aware of each other and “sympathetically participate” in each other’s lives for a given period of time. This presupposes that participants’ experiences are coloured by the immediate involvement—or the shared stream of consciousness—that they have with each other. An implication of immediate involvement is that the actor cannot stop to reflect on what is going on in the encounter without at the same time withdrawing his immediate involvement in it. The encounter has to be constantly and reciprocally experienced, meaning that it is lived through rather than observed. Therefore, it cannot in itself be grasped reflectively in the face-to-face situation.
The greater my awareness of the we-relationship, the less my involvement in it, and the less am I related to my partner. The more I reflect, the more my partner becomes transformed into a mere object of thought. (1932/1967:167)

The scenario presented by Schütz resembles Goffman’s discussion of joint spontaneous involvement\(^{30}\) (hereafter JSI). Goffman (1967) defines JSI as the reciprocal direction of attention to the conversation one is participating in, but also likens it to a ‘socialized trance’. In other words, JSI is depicted as a social phenomenon that resembles the idea of flow (see Csikszentmihalyi 1990), which gives it the character of a very intense and absorbed way of relating to the other and the encounter. Elsewhere in Goffman’s (1963:36) writings, involvement is defined as a form of engagement that is characterized by the sustainment of “cognitive and affective engrossment” in the activity and by the “mobilization of one’s psychobiological resources.” In either definition, the core of the matter is how actors allocate their attention, and here JSI comes across as a mode of relating, where the attention of participants is completely directed towards one single thing: the ongoing encounter. In less intense forms of involvement, actors can, however, simultaneously maintain both main involvements and side involvements:

A main involvement is one that absorbs the major part of an individual’s attention and interest, visibly forming the principal current determinant of his actions. A side involvement is an activity that an individual can carry on in an abstracted fashion without threatening or confusing simultaneous maintenance of a main involvement. (Goffman 1963:43)

One could perhaps say that JSI is a main involvement that does not allow actors to simultaneously engage in several, or at least not in very demanding, side involvements. It gets its spontaneous character when an individual who is already engaged in an activity becomes “caught up by it, carried away by it, engrossed in it-to be”, and when everything besides the activity appears irrelevant to the individual who is now an “integral part of the situation” (Goffman 1961:38). In adding spontaneity to the construct, Goffman emphasizes how it is impossible to control or plan involvement, or reflect upon it while one is involved. JSI must necessarily be non-reflected and unforced as every effort to make sense of or manage it will draw attention away from the social situation and thus reduce involvement. For the actor, then, there is a

---

\(^{30}\) The notion of involvement appears in several places in Goffman’s writings, and the indicated meaning of the term varies. In everyday speech, involvement is often supplanted with the more emotional-laden concepts “commitment” and “attachment”, writes Goffman, and elsewhere in his work, involvement is exclusively discussed in terms of engagement (1963:36, footnote 3).

Goffman (1963:36pp) at times describes involvement within a situation (situated involvement) as a purely inward feeling that is made visible through a learned type of body idiom called the involvement idiom. Involvement is also described as a capacity to give or withhold from giving one’s complete attention to some activity.
constant conflict between, on the one hand, ‘true’ involvement and, on the other hand, reflection and self-consciousness.

It is by now obvious that the notion of JSI shares much in common with Schütz’ ‘intense’ we-relationship: particularly with the idea that, when engaged in a we-relationship, the individual must put her own subjective stream of consciousness aside in order to engage in a common stream of consciousness. But as Goffman was not (unlike Schütz) very attentive to the experiential level of involvement, he has been accused of failing to “carry through the sociological implications of his more phenomenological insight that it is fundamental to human experience that we are always spontaneously involved in something” and that he thereby circumvents “the meaningful social world” (Ostrow 1996:346). According to Ostrow (1996:347), Goffman’s fear of retreating into psychologism rules out the possibility of him ever reaching a complete understanding of spontaneous involvement. By contrast, Smith (2005:408) asserts that Goffman’s work constitutes a development of phenomenological ideas as his analyses reach beyond the cogitating ego: “[h]is ensuing analysis of ‘self-consciousness,’ ‘interaction consciousness,’ and ‘other-consciousness’ might be regarded as an interactional phenomenology that grounds Schütz’ intellectualistic formulation of the we-relationship in the dynamics of interactional practices.”

These observations are reflected in the fact that Goffman pays more attention to the regulation of JSI than to its experiential nature. His main concern is how JSI and other involvements are subject to normative regulations, where participants take measures to follow involvement rules and obligations. The understanding of what is an appropriate degree of involvement in the encounter varies depending on how the encounter is framed, and in each encounter actors are made responsible not only for maintaining their own involvement, but also for helping others maintain theirs. Goffman (1967:115) puts much emphasis on involvement in his writings, so much so that he argues that the regulation of involvement is nothing less than the maintenance of social order. Since co-present individuals are always to some level involved in the situation, involvement furthermore occurs both in unfocused and focused interaction. However, the difference between the two forms of relating is, to Goffman, a difference in the degree of involvement, where actors are more involved in focused interaction. Following this, the intense involvement that JSI represents would only occur in focused interaction.

Distance Through Misinvolvement

As has been touched upon above, involvement in an encounter or a we-relationship is weakened in instances when attention is re-directed from social interaction to other objects or events. Thereby, it is possible to conceive of encounters where actors are physically co-located—that is, transmit ex-
pressed and embodied information—yet remain socially uninvolved and distant.

To Schütz (1932/1967:168pp), the we-relationship may not just be disintegrated by spatial or temporal distance, but also by reflective distance (i.e. distance through objectification, or to turn one’s interaction partner into an object of thought). Similar to the idea of reflective distance, Goffman (1963:69pp) proposes that an actor may be so preoccupied with reflecting upon past or future events that he fails to be involved in the present. That is, he will appear to be “away” and “not there”. Alternatively, a person can have “occult involvements” in which she is so deeply “away” that she is not even personally aware of her current condition. An extreme example of this is when a person suffers from hallucinations or a psychosis.

In “Alienation from Interaction” (1967), Goffman describes additional ways in which an actor may fall out of involvement and become alienated (or misinvolved). These types of disturbances are divided into four categories, namely external preoccupation; self-consciousness; interaction consciousness; and other-consciousness.

External preoccupation occurs when attention is caught by something unrelated to the current social interaction and its participants, while self-consciousness becomes problematic in the JSI context as soon as the individual is self-attentive in a way that disturbs her involvement in the interaction. But, unlike external preoccupation, self-consciousness is still “a kind of preoccupation with matters internal to the interactive social system” (1967:119), as it concerns how interactants withdraw from the interaction in order to reflect upon how other participants in the interaction see or judge them. A typical example is how a person, as soon as his definition of self is threatened, immediately stops being fully involved in the interaction itself, and instead becomes involved in thoughts about how he is faring in the interaction.

The two remaining forms of misinvolvement are interaction consciousness and other-consciousness. Like self-consciousness, they represent types of internal preoccupations, only they are not, in Goffman’s view, part of the interaction system in itself. Interaction consciousness occurs when a participant is so engaged in how the interaction is going, that she fails to involve herself in what the interaction is actually about. This may, for instance, happen to a dinner host who is so obsessed with the guests’ well-being that all effort goes to keeping everybody interactionally involved and the conversation pleasant and smooth sailing. While doing this, the host will not be able to fully engage in the ongoing interaction, and her behaviour will at best

---

31 Several of the points Goffman raises in “Alienation from Interaction” (1967) (where JSI is in focus) also correspond to what he proposes in Behavior in Public Places (1963). Only, in the former, the description of misinvolvement is somewhat more straightforward.
keep her alienated and at worst contaminate the very guests she is so eager to protect from alienation.

Other-consciousness occurs when one participant in interaction starts paying so much attention to other participants that he can no longer maintain involvement in the interaction. This typically happens when the behaviour of the other is deviant or improper, or when for some reason the appearance of the other attracts undue attention (e.g. an unfamiliar accent or a stigma such as a facial deformation or a speech defect). To take an example of improper behaviour, an individual may experience that the other is using the interaction only to promote her own interests in an insincere or affected manner.

Another example of improper behaviour is what Goffman (1967:122) calls “over-involvement”. In each situation there are limits to how involved in, absorbed by, and emotionally expressive persons are allowed to be in the ongoing interaction. Over-involved persons violate those rules by allowing themselves to become too emotional or too carried away by the interaction. Such lack of self-control may make the remaining participants allocate all their attention to the deviant behaviour of the over-involved person, with a consequential lack of involvement in the interaction. In other words, it is not the over-involved person who gets misinvolved, but the persons paying attention to her. Instead of focusing on what the over-involved other is saying, the misinvolved individual will focus on the other as if she were a mere object. Finally, the actor may also be “autoinvolved” in her own body (e.g. picking her teeth, putting on more make-up) in a way that both weakens her involvement in the social situation and attracts undue attention from other participants (and in effect causes them to become other-conscious).

Although it is theoretically possible to distinguish the above-mentioned forms of alienation from interaction, Goffman (1967) emphasizes that—in reality—they rarely make solitary appearances. When one type of misinvolvement occurs, a chain reaction of various misinvolvements is likely to follow, for example when autoinvolvement causes other-involvement.

**Intimacy**

Intimacy is the second mode of knowing the other that Schütz identifies as influential to the degree of closeness between actors. So what would intimacy as a way of relating entail? When previously discussing indirect social

---

32 Elsewhere, Goffman (1963:90) discusses several other ways in which behavioural deviance can disturb involvement. For instance, a person can attack the encounter on purpose or fail to follow the rules of the situation. Alternately, a lack of coordinated responses or a disruption of mutual commitment may disturb involvement; possibly because it creates a sort of situational contingency and insecurity.

33 In Goffman’s (1967:121) terminology, an affected individual is overly “concerned with controlling the evaluation an observer will make of them, and seem partly taken in by their own pose” while an insincere person is “concerned with controlling the impression the observer will form of their attitude towards certain things or persons, especially towards him, and seem not to be taken in by their own pose.”
experience, I mentioned Mjöberg’s (2011) idea of intimacy as a form of intense shared presence, and outlined one of the four modes of relating that intimacy encompasses in her model. While the already discussed mode (closeness) can occur in indirect social experience, the three remaining modes that Mjöberg proposes occur in direct social experience.

First, Mjöberg (2011:71pp) suggests that intimacy is characterized by a sense of boundlessness (“gränslöshet”), in which the actor experiences that the boundaries between himself, the other, and the relationship are diffuse or even non-existent. This could be described as a sense of being as one, where reflections about who I am in relation to others are absent.

If boundlessness is a sense of being as one, then complete directedness (“fullständig riktadhet”) is a sense of acting-as-one, in that it occurs when two individuals sense that their actions are in complete accord with the situation and each other’s actions, through a form of intense reciprocal attention between actors (2011:76pp). This can occur even when participants experience their own selves as clearly separate from each other’s. Complete directedness thus makes the situation in itself boundless, since all external factors that threaten to terminate or in any way alter the situation will immediately be either included or excluded from it, causing no threat to the situation in itself.

A third mode of relating concerns the experience of time during the meeting with the other. Mjöberg (2011:82pp) argues that while situations of intimacy are temporary—they have beginnings and endings—the actor’s experience of time in these moments solely revolves around the present moment (“nuflöde”). For actors to experience time like this, they cannot consciously reflect on it or in any way try to ‘capture the moment’, and so it acquires the character of a type of un-reflected acceptance of the flow of time.

These three modes of ‘intimate’ relating thus concern how actors relate to themselves, other participants, and the ongoing encounter. Together, they bear a strong resemblance to joint spontaneous involvement in that they represent a form of deep immersion in the situation—an immersion that is particularly vulnerable to any form of conscious reflection on what is going on. Reflections concerning who I am in relation to the other, what belongs and does not belong to this situation, and how I would wish to capture this moment, are all described as weakening the intimacy of the situation, and resemble what Schütz calls ‘reflective distance’ and Goffman calls ‘misinvolvement’.

Simmel’s (1908/1950) description of intimacy is a very different one. Intimate relating, to Simmel, is a way of knowing each other that all participants consider exclusive to this particular relationship. While intimacy can occur in larger groups, the sense of exclusivity is most prominent in the dyadic relationship. That is because both members of the dyad are equally important for the dyad to keep its form: if one person leaves, the dyad ceases to exist. By contrast, a group of three or more participants does not necessarily
disintegrate with the loss of one member, which means that the importance of each member (for the continuing existence of the group) is less than in the dyad.

While the ‘content’ accessed in this way of knowing usually concerns information that is thought to single out the uniqueness and particularity of participants and of the relationship, it is not the content in itself that determines whether the relating is intimate, but whether the sharing of the particular information is exclusive to the particular relationship. However, Simmel (1908/1950:126) argues that when this way of knowing the other happens, it must happen within a special type of relationship—one whose “whole affective structure is based on what each of the two participants share with only the one other person and nobody else”—in order to become intimate. As he explains:

[C]ertain external situations or moods may move us to make very personal statements and confessions, usually reserved for our closest friends only, to relatively strange people. But in such cases we nevertheless feel that this ‘intimate’ content does not yet make the relation an intimate one. For in its basic significance, the whole relation to these people is based only on its general, un-individual ingredients. That ‘intimate’ content, although we have perhaps never revealed it before and thus limit it entirely to this particular relationship, does nevertheless not become the basis of its form, and thus leaves it outside the sphere of intimacy. (Simmel 1908/1950:127)

In his discussion, Simmel is more concerned with capturing what an intimate relationship is, than he is with the activity of intimate relating, and here he seems to suggest that intimate relating, specifically intimate knowledge sharing, does not necessarily make the relationship an intimate one. In fact, he seems to suggest that actors must already have employed an intimate framework in the encounter in order to interpret the interaction within it as intimate.

Simmel’s idea of intimacy suggests that the way in which actors frame their relationship is important for how certain behaviours will work to draw them closer together. In a related essay, his analysis of “how factors of repulsion and distance work to create a form of being together, a form of union based on interaction” demonstrates how such frames can also maintain a distance between interactants (1908/1971:144). To Simmel, the state of being a “stranger” refers to a specific form of interaction or relating, where the stranger is at the same time near the other participants of an encounter and distant from them. As a form of relating, strangeness therefore happens in ongoing encounters, and not outside of them.

Characteristic of strangeness is a certain sense of objectivity or freedom in the encounter. This means that the mode of relating is not tied down by any social bonds between participants, but the approach the stranger takes to others (and others to him) has a more abstract character. The stranger is not
relating to others as particular individuals, and others are—likewise—not relating to the stranger as a particular individual.

The stranger is close to us insofar as we feel between him and ourselves similarities of nationality or social position, but he is far from us insofar as these similarities extend beyond him and us, and connect us only because they connect a great many people. (1908/1971:147)

However, Simmel emphasizes that this does not imply that the stranger is not fully involved or participating in the encounter—nor does it mean that the information that is shared is not personal or private. Instead, the distance created by strangeness results from the generalized way of relating, where the stranger comes across as someone who is rather indifferent to the particularities of the other and does not demand or expect much of the present or future relationship to him. One can potentially interpret this as a relating that therefore carries less risk than other types of relating, because there is no relationship of significance that can be damaged by strangeness.

According to Simmel, strangeness does not merely happen between people who do not know each other well, but can also occur when close friends or lovers realize that their particular relationships are not that exclusive or unique:

> It is strangeness caused by the fact that similarity, harmony, and closeness are accompanied by the feeling that they are actually not the exclusive property of this particular relation, but stem from a more general one—a relation that potentially includes us and an indeterminate number of others, and therefore prevents that relation which alone was experienced from having an inner and exclusive necessity. (Simmel 1908/1971:148)

Strangeness, therefore, is a mode of relating that can happen in any social encounter, between any two people regardless of their relationship history. If Simmel’s idea of intimacy is a mode of relating in which actors see each other as special and unique individuals and understand what is going on in the encounter as exclusive, then strangeness targets a form of relating where participants treat each other as general types and the relationship as interchangeable. Since it appears as if the behaviour in both modes can take identical forms—the disclosing of private information can for example occur in strangeness and intimacy alike—the difference between strangeness and intimacy seems to lie in how actors understand the encounter.

Summary and Discussion of the Microsociological Perspective

In the previous sections I have investigated what ‘being together’ could stand for in interactionist and phenomenological theories. First, I discussed how, in the literature, a ‘minimal’ encounter is described as standing on
three legs: Mutual recognition, mutual responsiveness, and some shared, or at least congruent, understandings of what is going on in the encounter. I also detailed two types of social meetings that could be described as ‘minimal’ encounters, in that they do not require much to emerge (or cease to exist). The one type (sociability) is characterized by activities that would seem to bring actors closer together, while the other (civil inattention) is characterized by activities that keep actors at a social distance.

Similar to some theories of social presence, awareness of other social beings (‘other-orientation’ in Schütz terminology) here stands at the core of any social encounter. However, for other-orientation to qualify as a social relationship it has to be reciprocated, meaning that a person who is being one-sidedly aware of another person is not in a relationship with him or her. In a related vein, Goffman suggests that a relationship of co-presence emerges when co-located actors have this mutual awareness of each other. In its loosest form, the behaviour of co-present individuals is unfocused, but as soon as actors begin to entertain a shared focus, a shared openness to communication, and pay closer attention to the acts and responses of themselves and others, then the behaviour of co-present individuals is focused. This is what Goffman calls an encounter, and the behaviour within such an encounter is very much shaped by the frameworks participants employ to understand what is going on. As exemplified by sociability and civil inattention, even minimally organized focused encounters can be very different in terms of how ‘close’ participants appear to be together in them, all depending on the frameworks in use. But unfocused interaction is also to some extent structured, and here it is reasonable to think of the employed frameworks as more general ones. To not enter a public environment naked is, with few exceptions, a widespread rule in Western societies, and this rule belongs to a rather general (and cultural) framework that actors employ in a wide range of situations, so as to not disrupt unfocused or focused interaction.

To put the discussion of minimal forms of social meetings in relation to online chat, it would seem that online chat more readily allows for focused than unfocused encounters. Provided that unfocused interaction occurs prior to any overt acts of communication, and that the “mutual awareness” involved in it requires that an actor is not just aware of another person, but is also aware that the other person is aware of her, actors may have greater difficulties establishing mutual reciprocity in this interactional stage. Of course, the actor can be made aware that another person is simultaneously logged into the online chat service, but this is only a weak indication that the other is in return aware of the actor. Instead, mutual awareness is perhaps more often established alongside overt opening attempts (e.g. sending and receiving the initial messages), which in Goffman’s terminology would mean that the interaction is already focused.

After having described how the literature depicts the ‘minimal’ requirements for actors to be together at all, I went on to examine how closeness
and distance is discussed in the literature. In line with Schütz (1932/1967),
the examination distinguishes between indirect and direct social experiences,
where indirect experiences could be said to be ‘distanced’ and direct experi-
ences ‘close’. The distinguishing factor here is immediacy: whenever two
actors cannot experience each other immediately, they are experiencing each
other indirectly, as contemporaries. While, according to the literature, indi-
rect social experiences would not suffice for any living social relationships,
co-presences, or encounters to occur, the discussion of contemporaries might
bear relevance to the new forms of near-immediate interaction afforded by
digitally mediated environments such as online chat environments.

Of particular interest is Schütz’ (1932/1967) description of how, even
among those who are contemporaries, some contemporaries appear to be
‘closer’ to the individual than others, because they are less anonymous to
him. What makes a characterological contemporary less abstract than, for
example, a habitual contemporary is that the actor has previously experi-
enced him or her directly. So while the characterological contemporary re-
fers to a unique individual to whom the actor has, or has had, a relationship,
other contemporaries refer to social types rather than particular beings.
Therefore, the difference in ‘closeness’ here comes down not to the degree
of immediacy, but to the knowledge an actor has of a contemporary and how
she has acquired that knowledge.

Mjöberg describes mental closeness as something that can occur between
an actor and her characterological contemporary. This is not an immediately
reciprocal relation between two particular individuals, but the relationship
between an individual and her mental image of another particular individual.
A relation of mental closeness could thereby be interpreted as a situation
where the actor’s stored knowledge and memories of past encounters with
the other form a mental (and probably also emotionally laden) representation
of the other, to which the actor in turn can relate. Of course, the sense of
closeness an actor would feel to such a mental representation would appear
to depend on what knowledge, emotions, and memories the representation is
made up of—relating to friends and loved ones in their spatiotemporal ab-
sence is perhaps more likely to induce a sense of closeness than relating to
one’s mental image of the check-out clerk (unless, of course, the check-out
clerk has made a strong impression). Mental closeness therefore concerns
how one specific individual relates to the particularities of another specific,
but spatiotemporally distant, individual (i.e. a characterological contem-
porary), and to the particularities of their specific past and present (and poten-
tially also projected future) relationship. For online chat, this would mean
that—provided actors are relating to each other as characterological contem-
poraries—they may feel a certain closeness to each other even when they are
engaging in text-to-text conversation at a spatial distance.

I thereafter turned to direct social experiences, and here I mostly dis-
cussed how the literature has described distance and closeness within fo-
cused interactions (encounters). While spatiotemporal immediacy can vary even within direct social experience, and thereby influence how close together actors are, the literature also shows how different modes of relating affect closeness and distance in encounters. The two types of relating discussed are involvement and intimacy. Involvement stands for a certain focused attention to the encounter, which in its more intense forms (for example joint spontaneous involvement) resembles something of complete and exclusive engrossment in the encounter. Involvement is disrupted when an actor fails to maintain her attention to the encounter, for example when she turns her attention inwards and engages in reflection, or when external objects or events distract her. In essence, the idea is that the less attentive actors are to the momentary flow of interaction, the more distant they become to each other and to the encounter. Spatiotemporal immediacy determines how much and with what frequency information reaches the actor, and therefore it also affects her attention to, and involvement in, a situation. The relatively low spatiotemporal immediacy of online chat would therefore suggest that it might be difficult for actors to reach heightened levels of joint involvement in these types of situations, which will be discussed further in the coming chapter.

While Mjöberg’s (2011) discussion of closeness in direct experience goes under the label of intimacy, I suggested that, together, the three modes of relating that she mentions portray a form of intense involvement, perhaps the most intensely involved actors can ever be in an encounter. By contrast, Simmel’s (1908/1950) idea of intimate relating is rather different from that of Mjöberg. Here, relating is intimate when actors understand what is happening in the encounter as exclusive to their particular relationship. That is why ‘strange’ relating—which Simmel (1908/1971) describes as a mode of relating in which actors do not quite approach each other as particular individuals—can to an observer look identical to intimate relating. At the same time, strangeness could be characterized as a distant form of togetherness while intimacy indicates a certain closeness between actors. The way in which Simmel’s intimacy and strangeness affect closeness or distance in an encounter, therefore, is through the common denominator of how participants understand the encounter, each other, and their relationship. Because of this, it would seem that intimacy and strangeness are not too dependent on the communication medium being used, meaning that the concepts are potentially applicable also in the context of online chat.

What is missing from Simmel’s discussion is, however, how actors arrive at an understanding of the encounter as intimate or strange. It is clear that this understanding cannot be formed on the basis of how actors behave (as the behaviour in intimacy and strangeness can be identical). Moreover, intimate relating can occur between people that have never before met, just as strange relating can occur between lovers. But Simmel does not discuss how people can ever tell the difference nor does he mention any situations in
which one person for instance misjudges another person’s strangeness as intimacy.

Assuming that the actor is personally and emotionally invested in an intimate relationship, while remaining rather indifferent in strangeness, one possible factor that could influence a person’s understanding is her perception of risk in the encounter—her willingness to put herself at risk of being rejected by the other, but also her trust in the other to not reject her when she is taking a risk. To take the example of the sharing of private thoughts and feelings, the actor appears to have more to lose, but perhaps also more to gain, by this behaviour in intimacy than in strangeness. Exposing parts of oneself that one normally conceals would—in intimacy—mean that the exclusivity of the relationship (and thus the intimacy) increases; at the same time, any potential rejection would become a matter of an increasingly unpleasant rejection of one’s (perceived) ‘true’ self. By contrast, the detached mode of relating characteristic of strangeness suggests that the actor is more indifferent to what the other thinks of her personally, meaning that even if she exposes herself in the encounter and is rejected because of it, this rejection should be less painful. If strangeness is a mode of relating where actors are more indifferent to rejection, then strange encounters would involve fewer personal risks than intimate encounters. An intimate mode of relating, by contrast, would be a mode in which the actor truly cares about what the other thinks of her as a unique individual. Intimacy and strangeness, therefore, could perhaps be described as attitudes of caring or indifference—attitudes that would depend on the meaning (or understanding) actors ascribe to the encounter.

The discussion of involvement and intimacy suggests that the more intimate and involved actors are in an encounter, the more close together the actors can be said to be, while the distance between them increases with the occurrence of strangeness and misinvolvement. As the two concepts imply different types of closeness or distance, actors can of course also be misinvolved in intimacy, or ‘strangely’ involved. In addition, Asplund’s notion of asocial non-responsivity similarly suggests that, whether it is voluntary or involuntary, actors may at times completely cease to be responsive to each other, even when stimuli are available. So overall, the literature strongly points to social absence (as represented by asocial non-responsivity, misinvolvement, or strangeness) as something that may occur in direct social experience, which begs the question of whether such social absence can also happen in online chat and other near-immediate interactions. Along similar lines, the notion of mental closeness points to the possibility that spatiotemporally distant actors can still experience a sense of closeness to one another, for example while engaging in an online chat conversation.

It is easy to perceive distance in interaction as unwanted and largely accidental, and as either a spatiotemporal distance that is out of the actor’s (momentary) control, or as a form of failure to be intimate or to maintain focus
on the encounter. But if we stop at that, we miss that distancing behaviour in interaction is often both voluntary and socially expected. Asplund pointed to this in describing how the free and spontaneous social responsivity we are born with is moulded into a much more discerning, disciplined and socially acceptable responsivity as we grow older. Similarly, the meeting between two strangers does not just result in low-involved behaviour (like civil inattention instead of sociability) by chance, just as the reason a meeting between a bank employee and a client is usually characterized by strangeness is not that the actors have failed to be as intimate as they would be on a romantic date. What must also be added to the equation is how actors relate to others in accordance with their understanding of the situation, and the overall behavioural norms governing it.

In his analysis of involvement in focused and unfocused interaction, Goffman (1963) is very clear about this. His main concern is how involvement is subject to normative regulations, and he argues that co-located actors are always under the obligation to follow involvement rules. According to Goffman (1967), the highest level of joint spontaneous involvement is rarely wanted in a social situation, and most situations rather require a certain amount of misinvolvement on the part of their interactants. This demand is often met by a deliberate disruption of involvement—so-called spacing behaviour—by the utilization of “involvement shields” that block perception. The individual thus makes it impossible for others to observe her, although she may arrange it so that she herself can still observe others.

So, in Goffman’s theory, behavioural norms and roles often function so as to limit or shape involvement and thereby they ensure a certain distance between actors. But deviances from the behavioural norms can also cause unwanted misinvolvement in a situation. When an actor, for instance, acts in a way that conflicts with the shared definition of the situation, she is likely to disrupt other participants’ involvement. That is because their attention to the deviant behaviour will now distract them from pursuing their own involvement obligations, and so the deviant, and therefore distracting, behaviour has the effect of making other participants misinvolved.

As for the way social organization affects intimate modes of relating, these modes appear to be rather tied to informal versus formal roles. While intimate relating is rarely advised in formal encounters in which actors take on professional roles (such as in the example of the bank employee and the client mentioned above), strangeness is typically avoided in informal encounters, particularly when actors employ frameworks that suggest that the relationship to the other is defined as ‘close’ (as for instance a friendship or love relation is). This can be illustrated further by Goffman’s (1959) distinct

34 Spacing behaviour is the kind of behaviour that ensures that a proper distance between two or more individuals is maintained (Goffman 1963:162).
tion between front and back region behaviour. In front regions, actors are expected to maintain a distance to the back region behaviour by limiting how intimately, personally, and also spontaneously they act, while back region behaviour must be distanced from excessively impersonal, business-like and scripted behaviour.

Social behaviour in digitally mediated environments is of course also subject to normative regulations that vary across environments. There are digital environments that are more public (e.g. many online chat rooms), more private (e.g. most instant messaging environments), more formal (e.g. a company’s instant messaging service), and more informal in nature (again, many online chat rooms). Each of these ‘types’ of online environments of course have different rules for what is considered appropriate behaviour, for example in terms of how intimate and involved one should be (and how to express this involvement and intimacy), even though the means of information transmission may remain largely the same.

In a public chat room, there is usually no stated rule about who can talk to whom, as these environments are often created with the intention of allowing people to meet new people. Yet certain behaviours might be discouraged. Some content (e.g. sexual or racist), flaming behaviour (i.e. sending hateful messages), spamming (i.e. sending a large amount of the same message, or nonsensical messages, over a short period of time) can be prohibited in the chat room and cause the deviant actor to be permanently or temporarily exiled from the environment. By contrast, instant messaging is usually not moderated by an external party, and instead the involved actors have more freedom to themselves negotiate the ‘rules’ of the conversation. As users have to be added to each other’s contact lists prior to initiating interaction, this is also not an environment in which actors normally ‘meet’ each other for the first time.

In summary, the degree of closeness and distance in face-to-face encounters cannot merely be explained by environmental circumstances, nor can the behaviours actors engage in to increase or reduce distance in encounters be depicted as just happening by chance or completely under the control of the actors. The influence that frames have on behaviour that increases or decreases distance in encounters cannot be overlooked, and it is reasonable to think that the same goes for meetings in digitally mediated environments.

Discussion and Summary

In this chapter I have reviewed two discourses that deal with the phenomenon of being there together. To begin with, I examined the literature on social presence, which is a concept used in the study of digitally mediated social interaction. Thereafter, I turned to the study of togetherness in physical
environments, where I particularly focused on interactionist and phenomenological studies.

Broadly speaking, ‘being there together’ represents three different things in social presence theory: an affordance of the medium, an experience resulting from the actor perceiving various events in the environment, and as related to self-presentation, and I discussed what each of these definitions would imply for social presence in online chat. While many descriptions of social presence frame the phenomenon as one that only relates to electronically and digitally mediated environments, there are exceptions. For example, both Biocca and colleagues (2002; 2003) and Riva (2008a) suggest that the experience of ‘being together’ is a matter of the actor’s awareness of other people in any environment.

A concern with the actor’s awareness of others, or how the medium can afford such awareness, is a common theme in social presence theory. That such an awareness is necessary for togetherness is also acknowledged in the microsociological literature, for example when Schütz (1932/1967) argues that being in the social world assumes that the actor is other-oriented. But to Schütz, other-orientation only qualifies as a social relationship when it is reciprocated, and Goffman (1963) similarly states that co-presence represents a form of mutual awareness—i.e. unfocused interaction—between actors. As co-present actors begin to explicitly cooperate to sustain a shared focus of attention, typically through communicated information, they enter into an encounter with each other. Here, actors must be open and responsive to communication, share the ambition to engage in focused interaction, and also behave in accordance with the framework applied to the situation.

With the exception of when social presence is defined as a form of self-presentation, the main difference between theories of social presence and microsociological understandings of togetherness is how the actor is portrayed. While, in social presence theory, the actor is often depicted as little more than a receptor of stimuli, the microsociological actor is someone who takes an active role in how she is being together with others, for example by negotiating and co-constructing the frames employed in the situation.

What both research fields have in common is the realization that distance matters when it comes to our being together with others. However, while social presence research tries to understand how people can be together despite spatiotemporal distance; the possibility of spatiotemporally distant togetherness is for the most part unthinkable to the classical microsociologists. For instance, Goffman (1963:22) states that mediated interactions, like telephone calls or letters, are ‘merely’ situated, while the full conditions of properly situated co-presence can only be met in direct social experience. Mediated interactions, therefore, can never become focused in Goffman’s understanding, meaning that, from his perspective, actors cannot encounter one another in the digitally mediated environment.
Although this stance can perhaps rightfully be excused by the fact that, at the time it was taken, digitally mediated interaction had yet to see the light of day, its exclusive focus on face-to-face interaction makes it difficult to directly apply to digitally mediated interactions like online chat conversations. But, perhaps as a result, the cited microsociologists also do not get caught up in thinking that external, environmental factors (such as the medium’s capacity) are the only things that matter for togetherness. Although the spatiotemporal conditions of face-to-face encounters are still considered influential (as illustrated by Schütz’ discussion of immediacy), closeness is also thought to be determined by factors that are not directly related to the environment. For example, aspects internal to interaction, such as intimacy and involvement, are identified as just as likely to affect the degree of togetherness. These modes of relating are, in turn, shaped by how actors frame the encounter, meaning they are governed by behavioural and social norms and structures. Since online chat environments and other digitally mediated environments are also populated by people operating in accordance with different frames and social structures, the microsociological concepts may indeed be useful for the analysis of both digitally mediated and face-to-face togetherness. This is not to ignore that some social presence scholars also point to interactional elements as either causing or influencing actors’ togetherness, but in most cases these factors are not explored in any greater detail.

In conclusion, social presence theory and classical sociology seemingly constitute two quite contradictory approaches to the phenomenon of togetherness. In the first camp, it is assumed that it is possible to be there together at a spatiotemporal distance, and in the second, it is assumed that such circumstances render togetherness unlikely. Because of their focus on spatiotemporally immediate interaction, the microsociological theories are not in themselves sufficient to understand togetherness in digitally mediated environments, but instead they have the advantage of taking seriously how, in accordance with behavioural norms and structures, actors themselves contribute to creating or diminishing distance in encounters. Social presence theory, by contrast, puts most emphasis on how the digital medium allows actors to be together, paying less attention to how interactional elements shape togetherness. While I have already pointed to ways in which the concepts described in this chapter may be useful (or not so useful) in the analysis of togetherness in online chat environments, in the next chapter I will examine this relationship in more detail.

As the preceding chapters have shown, the phrase ‘being somewhere together’ can mean a variety of things depending both on what we ascribe to the word ‘somewhere’ and to the phrase ‘being together’. The problem of modern day togetherness, it was argued, originates from the changing nature of the *where*s in which we can be situated. This is in turn an effect of how the environment now affords us the opportunity to reach out to and interact with people that are not in our spatiotemporally immediate presence. Put another way, the contexts in which actors can meet are no longer limited to shared physical situations, but extend to digitally mediated *where*s of being together.

However, there are still many different ways in which one can conceptualize this *where*, and in this chapter I will argue that the main difference between social presence theories and microsociological theories about togetherness lies in how the *where* of our togetherness is understood. In addition, this chapter will employ the example of online chat services to examine which approach to togetherness appears to have the greatest analytical value. The chapter is divided into two sections of which the first deals with the *where* of social presence theory, and the second deals with the *where* of microsociology. Before that, however, I will briefly expand on the categorization proposed in Chapter 2, namely that which broadly distinguishes between three *where*s in which actors can be said to be together: the informational environment, the larger social situation, and the specified social situation.

Given the last chapter’s discussion I now believe this categorization can be made more specific, particularly with respect to what being together would stand for in each of the categories. Based on what is described to support togetherness, togetherness can firstly be portrayed as the co-situatedness (hereafter: co-location) of actors in an informational environment (which is physical, or digitally mediated). Togetherness here primarily concerns the informational access actors have to one another, which is a kind of shared set of affordances of information flow. However, while co-located actors may have the possibility to access information about each other, they do not need to stand in any form of reciprocal relation to each other. In other words, they
do not currently need to stand in any mutual relationship of awareness, attention, or access. For example, two individuals can be co-located on a public street or in an online chat room (e.g. by being simultaneously logged in to the chat room), yet have no established relationship of reciprocity going on between them. What is important to co-location is first and foremost how co-location is shaped by the informational affordances of the environment, and not the way in which the actors relate (or do not relate) to each other. Included in this category are many (if not most) of the reviewed definitions of social presence, but Schütz’ non-reciprocal other-orientation would also belong here.

Second, co-located actors can be co-present in an (unfocused) social situation, meaning their relationship to each other is structured not just by the informational affordances of the environment, but also by a larger behavioural structure (or frame). In co-presence, actors are reciprocally accessible, available and subjected to one another. However, while participants in unfocused situations will apply a framework to the situation, and adjust their behaviour accordingly, these adjustments still do not need to occur as direct and overt responses to the presence of the other. Take for instance a person who is sitting alone in the workplace cafeteria when another person walks in. Even if no greetings or overt communications occur between the two actors, they will be reciprocally aware of each other and they will now be less likely to engage in behaviours that are considered inappropriate (such as nose-picking or other autoinvolvements). This type of unfocused awareness and sense of mutual monitoring also occurs in digitally mediated situations. A person who makes a status update on Facebook for all of his 435 Facebook friends to see might not compose this message with anyone in particular in mind. Yet he is aware that his actions will be noted and monitored by others and this awareness will be reflected in his conduct (e.g. that he writes a post at all, and what he decides to write in it). If he, for example, was the only person on Facebook, or if he knew that no other person on Facebook could see his status updates, he might not have made an update in the first place. Around the web one can find similar examples of behaviours that serve to make the actor visible to everyone yet no one in particular, for instance in certain public updates on weblogs, photo-blogs (e.g. Instagram), video blogs (e.g. Vine), or microblogs (e.g. Twitter).

There are, of course, some significant differences between the two examples mentioned here. In digitally mediated situations, co-presence does not (as Goffman would have it) mean that actors can access each other with their naked senses; they are not in each other’s immediate presences. There is therefore a difference between how motivated actors must be to make themselves co-present in physical situations compared to digitally mediated ones. While the actor who is sitting in the cafeteria may have no wish to enter into a relationship of unfocused interaction with the actor who is walking in, and vice versa (their respective goals may solely be to finish their lunches and
get back to work), the actors who are posting or viewing others’ updates on
social networking sites must have some form of social agenda. One does not
simply ‘stumble’ into unfocused co-presence in the digitally mediated envi-
ronment, as one can do in physical environments. But at the same time as
actors can assume that co-present others in an online chat room are there
because they in fact want to be there, it may be harder to know for certain
that those seemingly co-present others are actually co-present without enter-
ing into a more focused encounter. For example, as Schütz (1932/1967:203)
points out, the indirectness of the relationship between contemporaries al-
ways involves this element of chance: one can never know that someone else
is truly ‘there’ in for example the online chat room until someone makes an
overt attempt at actually finding out.

All these things considered, the category of unfocused social situations
would comprise the conceptualizations of togetherness that portray the phe-
nomenon as involving some form of mutual awareness or observation, such
as Schütz’ living social relationship, Goffman’s notion of unfocused co-
presence, and also some descriptions of social presence that portray the phe-
nomenon as a relationship of reciprocity or self-presentation.

Third, actors can be co-present in focused social situations. Focused co-
presence would be structured both by the informational environment, the
larger social frames present in unfocused situations, and by more specific
frames. In these encounters, actors must be sustaining an overt and mutual
connection to each other. They are reciprocally aware, attentive, and respon-
sive to each other and they have a shared understanding (or congruent unde-
rstandings) of what is going on. In a physically defined informational envi-
ronment, focused co-presence can, for example, be represented by a conver-
sation or by an exchange of explicit bodily signals (e.g. different gestures),
and in a digitally mediated informational environment it can be represented
by an exchange of written messages, a video conversation, or a phone call.
This category would include Goffman’s notion of focused encounters, and
some definitions of social presence as self-presentation.

As shown, togetherness in digitally mediated contexts can either be stud-
ied from the perspective of co-location, or the perspectives of unfocused or
focused co-presence. It follows that the particular where scholars have in
mind when they discuss being somewhere together is bound to influence
their understandings of the phenomenon in question. Zhao and Elesh (2008)
have made a similar observation in distinguishing between two ‘spheres’ of
social contact: the sphere of co-location and the sphere of co-presence. Co-
location, they argue, is a spatial relationship that puts people within range of
each other, while co-presence is a social relationship that puts people within
mutual accessibility for contact. However, Zhao and Elesh do not distinguish
between unfocused and focused social situations, and are mostly concerned
with the norm-driven regionalization of social contact in digitally mediated
contexts. They point out that the fact that Internet technologies render actors
within range of (practically) all other Internet users does not mean that all users are, in fact, within reach of each other, and so electronic access does not equal social access. In other words, co-presence is here discussed from the perspective of social inclusion and exclusion, where factors other than the technological ones decide who can contact whom, and who will interact with whom. An actor’s access to situations of co-presence thus varies from region to region, and it is influenced by her status and other actors’ willingness to initiate contact. While I do not disagree with Zhao and Elesh, how social status and norms for inclusion and exclusion influence access to social situations, and thus togetherness, falls outside of the scope of this thesis.

The following two sections will, based on the literature reviewed in Chapter 3, discuss the implications of applying the perspectives of co-location and co-presence (unfocused and focused) to online chat. I will firstly discuss social presence theory, and argue that the favoured view in this field is limited. That is because its focus on co-location (and occasionally unfocused co-presence) is too narrow to fully capture what it means to be together, let alone how actors get closer together, in online chat. Thereafter I discuss the microsociological perspective of togetherness, and here I particularly emphasize more focused forms of social contact. This literature is limited in the sense that it only considers co-situatedness in physical environments, and therefore I discuss if, and to what extent, the microsociological understanding of focused togetherness can be useful for the study of togetherness in online chat (and by extension, digitally mediated togetherness).

Social Presence Theory and Co-located Togetherness

To briefly recapitulate the findings from the review of social presence theory, the two main approaches to social presence are to either describe the phenomenon as an environmental affordance, or as an internal reaction to the outer world—an experience of things, social beings, and/or happenings in the external environment. A third and outlying position to social presence describes the phenomenon as a form of self-presentation. This last understanding bears more relevance to the coming discussion of co-presence, and will therefore not be included in the immediate discussion.

An alternative approach to the literature would have been to categorize social presence in terms of its application. Here, one branch sees social presence as social perception, and touches both upon how actors must be perceivable in the environment as well as on the psychological reactions that social perception triggers. Even though digitally mediated encounters are in focus, the perspective is general enough to be applied to all kinds of social encounters: mediated encounters and face-to-face encounters alike. The second branch speaks of social presence as something unique to mediated encounters, most notably in terms of a replication of the ‘real’ social presence that
happens in face-to-face environments. This branch is therefore interested in what elements of face-to-face meetings are necessary to replicate in digitally mediated environments in order to make these digital meetings ‘real’.

Drawing on the discussion in Chapter 3, one can also conclude that what the reviewed definitions of digitally mediated social presence share is that they all see the where of being somewhere together as an environmentally defined situation or in some rare cases a loosely organized social situation (for instance when social presence is described as a form of reciprocated awareness). But even when social presence could be said to stand for a form of co-presence, the main emphasis lies on how the phenomenon is shaped by the environment, and not on how it is formed and shaped in the situation. Thereby, social presence theory is a theory that is first and foremost concerned with actors’ co-location in environments; with how environments allow people to detect other co-located actors, what digitally mediated co-location is, and what digitally mediated environments must afford for co-location to happen.

However, within the field, there are some differences with regard to how this environment is depicted. While some scholars assign social presence to any environment that can support perception of others, others portray the environment of social presence as one that is clearly distinct from physical environments. In this latter perspective, we also see signs of what Graham (1998) warned about, namely that the digitally mediated environment is being portrayed as a spatially non-spatial environment that will always be inferior to the ‘real’ spatial environments of the material world. For example, many scholars argue that digitally mediated social presence only comes into being when actors are tricked into believing that they are in a face-to-face meeting, and so digitally mediated social presence is depicted as an illusion of non-mediated, ‘genuine’ social presence. Digitally mediated social presence will, in this perspective, always have an ‘as if’ (it was non-mediated) hanging over itself, because actors simply cannot be co-located in digitally mediated environments like they can be co-located in physical environments. Arguing to the contrary would be to adhere to the laws of science fiction, where people’s bodies are sometimes, quite literally, transferred into the spatial virtual world at the expense of losing their grip on the non-virtual world (see e.g. Tron: Legacy 2010).

Although the literature on social presence is not primarily concerned with social situations, many scholars do acknowledge that situational aspects, such as culture and norms, may influence social presence. However, these factors are comparatively unexplored. There might of course be practical reasons for this. As it is easier to improve the technology that supports an environment than it is to tamper with all the combined, and sometimes rather ungovernable, factors that make up a durable (and enjoyable) social situation, co-location should also be easier to engineer, and perhaps find more universal solutions for, than co-presence is. In the following sections, I will
nonetheless discuss how well the perspectives of co-located togetherness manage to capture actors’ being together in digitally mediated environments, and I will throughout make use of the example of online chat to point out some limitations in these approaches.

Co-located Togetherness and Online Chat Environments

Co-located togetherness predominantly concerns the actor’s detection and awareness of others in the environment, and so becomes a study of social perception and the nature of the perceived social cues. The field of social presence research is particularly focused on the psychological reaction to perceiving others in the environment, and to the affordances that will make possible the desired social presence ‘response’. So even though these theories concern the existence of two or more social beings in an environment, they are not primarily concerned with the interactions between those beings or with the frames shaping them.

There are, as I have pointed out, two different understandings of co-location in the social presence literature, where the one (the ‘universal’ approach) would encompass co-location in all types of environments, and the second only applies to co-location in mediated environments. As the second approach describes social presence in mediated environments as contrasting social presence in face-to-face environments, the two approaches are to some extent contradictory. I will begin the discussion by questioning an assumption made in the ‘universal’ approach, and then continue to question the main assumption of the ‘media-specific’ view. Thereafter, I will discuss problems that pertain to both of the standpoints. Throughout, I will exemplify my points by drawing on the case of text-to-text interactions (i.e. online chat interactions) occurring both between humans and between humans and computers.

Digitally Mediated Environments do not challenge Basic Social ‘Instincts’

The ‘universal’ approach to social presence describes the phenomenon as the psychological process of detecting and observing other social beings in the environment (Riva 2008a; Biocca 1998). This is described as a largely automatic process, and it is also seen as a crucial precursor to social interaction and intersubjective understandings of others. The approach bears a resemblance to the microsociological portrayal of a social instinct that, by probing them to react socially to stimuli that come their way, leads humans to engage in sociable activity with others—sometimes even in cases where the stimuli do not originate from another social being (e.g. Simmel 1910/1971; Cooley 1902; Asplund 1987).

But for modern day social interactions like online chat interactions, what has changed is not the human impulse to react socially to stimuli—what has
changed is the environmental conditions structuring behaviour, and what stimuli our environments can afford. And, as many studies have shown, the kind of stimuli an environment must afford to trigger the ‘social reaction’ in humans do not by any means need to be rich, ‘face-to-face’-like, or genuinely social.

Through a number of different psychological experiments, Reeves and Nass (1998), found that the social cues available in a given environment do not need to be numerous or very face-to-face-like for the actor to believe that she is encountering another social being. In fact, even very rudimentary social cues proved to be sufficient for that purpose. In addition, without explicitly being informed of an actor’s artificiality or of being aware of the possibility that some actors in the environment could be artificial, the human subjects would—by default—classify others in the environment as human others (Reeves & Nass 1998).

Humans also have a well-documented tendency to react socially to non-humans, even when they are fully aware of their non-humaness—for example, to objects such as robots and computer programs, or living, nonhuman, beings such as animals. When it comes to technology, Ihde (1990:107) describes how humans sometimes come to experience technological things (like computers) as ‘quasi-others’. In these ‘alterity relations’ the technology is no longer perceived as a thing or a useful tool, but rather as an ‘other’ that the actor can relate to. As we will see later on, so-called chat bots are known to (at least initially) entice the interest of humans in online chat, which suggests that they are more than capable of triggering the social instinct in humans.

The ‘social impulse’ is therefore not difficult to evoke in environments where there exists actual or perceived evidence of other social beings. It is not obvious how a digitally mediated environment like an online chat room would challenge or alter the human predisposition to react socially to stimuli. As online chat technology is a means for interaction that is also being widely used, it is reasonable to think that this technology can successfully help trigger the impulse to be socially responsive. In terms of affordances, there are of course many differences between different communication environments, but they all have in common that they afford the transmission and reception of social stimuli. Therefore, one could go so far as to say that all socially interactive media, at their very minimum, can trigger the social impulse in humans, even if they do not afford actors the opportunity to form more complex social bonds.

---

35 For an actor to go from interacting with a thing as the object/nonhuman being it is, to engaging socially with it, she must, however, consider it as being in possession of a social self. But according to Cohen (1989), the tendency to in this way anthropomorphize non-humans is very common in humans.
Let me illustrate this with two examples. In a shared physical environment, I may sit facing my friend while conversing with her over dinner, or we may keep a conversation going while being located in two different, visually cut off, rooms of the house. In both cases, the social cues I apprehend are equally good at telling me that my friend is within perceptual reach, and they may be equally good at prompting me to in some way react to her socially. Thus, the co-locatedness of my friend is more or less stable, even though fewer perceptual cues are used in the second scenario (only vocal, as opposed to visual, vocal, and perhaps even olfactory and haptic, in the first scenario). In an online chat scenario, I may be exchanging text messages with my friend over an instant message service, and this exchange can be equally efficient at establishing that I have perceptual access to her, as are having her in my physical proximity, or hearing her in a telephone call. The notification sound that my phone gives off when a new text message arrives in my WhatsApp is furthermore just as good as the sound of an incoming phone call is at triggering in me an impulse to pick up the phone and respond.

Of course, the more perceptual channels are put to use, the more perceptually ‘flexible’ can actors be. When I am sitting at the table with my friend, neither of us will need to speak to confirm that we are together in the room, because we can rely on our visual and perhaps also tactile and olfactory perceptions of one another. We could equally also be blind folded and not close enough to touch, but still establish mutual awareness through the use of speech, and possibly our olfactory senses. In online chat, on the other hand, the perceptual range is typically limited to visual perception (and the occasional auditory perception of sounds, for instance the one that notifies that the actor has received a new message), so there are fewer ways to establish that the other is perceptually accessible. However, in all the mentioned examples, each environment at the very least affords actors the ability to perceive and be perceived by others.

How the environment allows us to recognize other social beings does, however, introduce new elements of doubt. In the case of an online chat room, the actor is made perceptible not by his physical body but by his ‘symbolic’ body—in this case normally a user name and the occasional profile picture or symbol—and from that initial information it is impossible to draw any conclusions about, for example, what the person behind the symbolic body is directing his attention to and whether it is at all possible to reach him through his symbolic body at this point in time. It is often the case that an actor’s symbolic body remains perceptible to others in the online chat room after the actor himself has withdrawn from the environment, for instance when the actor leaves his computer to go to the bathroom. This simply does not happen in shared physical environments: an actor cannot continue to express co-location in a shared physical environment without bringing his body, and the cues that it automatically expresses, with him.
To return to the symbolic body, the fact that it is perceptible does not automatically mean that the owner of the body is capable of engaging in reciprocal social behaviour. This is true for physically co-located actors as well, where physical proximity does not necessarily mean that actors are socially accessible (i.e. ready and able to enter into a relationship of co-presence) to each other. But while physically co-located actors can quickly conclude how socially available the other is based on the expressions her body gives off, a symbolic body does not automatically give away such information. This makes it harder for actors to determine if what they perceive as digitally co-located others are—in this very moment—also socially available others. Many online chat services do, however, offer a solution to this problem, either by automatically informing users of how active others are in the environment (e.g. by indicating if they are ‘away’ or ‘active’), or by letting the actor himself label his current level of availability (e.g. ‘busy’, ‘online’, ‘working’, ‘be right back’). In that sense, online chat environments have been appropriated so as to ensure that the automatic assumption that co-location indicates a simultaneous social accessibility holds as true as possible. Despite those appropriations, and as mentioned earlier in reference to Schütz, actors cannot, prior to making an overt attempt at initiating interaction in the online chat environment, reach the same level of certainty (as they can in a physical environment) about whether others are indeed co-located, or better yet co-present, with them. However, as social presence theory remains focused on the social processes of detecting co-located others prior to interaction, it risks overlooking that actors can only reassure themselves that others are co-located in an online chat environment (and other digitally mediated environments) after overt communication is initiated.

Examining these basic social ‘instincts’—or elementary social perception—could, of course, be a relevant quest in itself. But for a field that is set out to broaden the understanding of digitally mediated social life and that claims it is doing so by investigating a phenomenon responsible for a range of desired, and quite complex, social outcomes, the just mentioned approach appears to be slightly misdirected. Since it only denotes a non-reciprocated detection of others in the environment, the low-level social presence described by Biocca and Harms (2002), would for instance not suffice for the opening of a social act (Miller et al. 1975), the emergence of a living social relationship (Schütz 1932/1967) or co-presence (Goffman 1963). Meanwhile, Biocca and Harm’s (2002) idea of high-level social presence encompasses the three steps required for an opening in Miller and colleagues’ theory, but they fail to take the definition further, so as to include anything in interaction that happens beyond reciprocal social awareness.

Non-reciprocated and reciprocated social awareness cannot, for example, account for how people become more deeply entangled with each other. That is, it does not account for anything but the comparatively shallow state of perceived or actual mutual recognition or awareness, which would be part of
any possible encounter: disinterested glances between people on the subway platform, a flirtatious first meeting, an angry (one-sided) conversation with a malfunctioning computer, a reunion of old friends, a user noticing that another user has come online in the chat room, an intriguing business conference, a dog owner playing fetch with her Golden Retriever, an uninspired lecture, a polite but distant encounter with the check-out clerk, and so on. What character all these latter stages of encounters take is also quite possibly more dependent on the frame and how actors choose to act in it, than it is the environment. The duration of an encounter, the interest actors have in it, and the meaning ascribed to it, are of course ultimately dependent on actors being able to recognize each other in the environment, but the recognition in itself only plays a small part in what happens next.

That is not to say that ‘what happens next’ must always denote a state of endearing closeness between actors, as exemplified by Goffman’s (1971) notion of civil inattention. Here, the requirements of an opening are met, yet the nature of the interaction still has the character of a very distanced and—might I say it—untogether-like encounter. Though social presence would appear to be a more complex social phenomenon than civil inattention, the social presence described by Biocca & Harms (2002) and Zhao & Elesh (2008) remains at the level of reciprocal attention, while civil inattention involves mutual responsiveness and congruent functional goals. That is because civil inattention requires actors to coordinate their behaviour with the shared goal of avoiding further contact. Otherwise expressed: they enter into the stages of an opening to ensure that they do not proceed into the next phase of social interaction (the middle). Of course, in this example, the line between co-location and co-presence may come across as very fine, yet I would argue that there is a clear difference between the kind of incapacitated observers described in social presence theory, and the civilly inattentive actor who conducts herself in accordance with situational norms as well as with her goals and intentions.

‘Real’ versus ‘Virtual’ Co-Location

The media-specific understanding of social presence depicts the phenomenon as a special type of co-location that is unique to mediated environments. At the same time it is thought that the more this environment resembles a face-to-face environment, the better the mediated co-location. The argument goes that different mediated environments are unequally equipped to afford (mediated) social presence, while the genuine form of social presence is guaranteed to be afforded in all physical environments.

As a general rule, the social cues transmitted in face-to-face communication are assumed to be best at evoking social presence, and, therefore, the quality of cues afforded by electronic media are compared with the affordances of the face-to-face environment. In effect, several researchers emphasize the difference between the face-to-face medium and various techno-
logical/digital social media (e.g. Short et al. 1976; Daft & Lengel 1984; Cullen & Markus 1987; Rice 1993). These theorists define social presence as a quality of the social medium, and assume that different media will convey qualitatively different social cues. For example, proponents of the so-called “cues-filtered out” perspective argue that the “social presence capacity” of a medium is decided precisely on the grounds of how many, how qualitative, and how informationally rich cues the medium can transmit. In other words, the cues-filtered out perspective maintains that the more technological social media resemble the face-to-face medium, the better for social presence. As I have earlier stated, the social presence potentials of online chat technologies do not look good through this lens. As it mainly relies on text, the non-verbal cues of the medium are minimal and thus its richness is considered very low. In addition, the rate at which information can be transmitted is also low compared to face-to-face interaction, telephone calls, and video calls, which further undermines the medium’s social presence ‘score’.

The idea that the medium affects how, and what types of, social cues appear in an environment is of course correct (on the verge of being a truism). Moreover, the face-to-face environment remains the environment that affords the most spatiotemporally immediate transmission of cues, and it is also the environment that affords the widest perceptual range. But, despite these affordances of the face-to-face environment, it is not entirely obvious that the most ‘natural’ form of interaction (i.e. face-to-face interaction) must necessarily be better at inducing togetherness than an online chat environment is.

In the context of human-computer interaction design, O’Hara and colleagues (2012) have made a similar observation. According to them, there is a problematic tendency to strive to come up with designs that allow for as ‘natural’ interaction as possible, where ‘natural’ stands for an interaction that is close—or ideally identical—to face-to-face interaction. Furthermore, this particular problem is portrayed as a representational, rather than interactional, one.

[T]hese interactions are seen as having an ideal, static and definable state and, though they are not always completely clear or exactly represented in any particular instance, they are something that can be, with sufficient understanding and scientific research, represented and modeled. (2012:5)

Inspired by phenomenology, O’Hara and colleagues go on to suggest that the naturalness of an interaction is nothing but an “occasioned property.” By that the authors mean that it is something that is “produced and managed together by people in particular places” (2012:5) and that allows actors to act in ways that fit the social circumstances they are in. What really matters to actors, then, is not how well interaction in an online chat room represents or mimics face-to-face interaction, but whether or not the medium is capable of meet-
ing their interactional needs. This brings us back to the difference between approaches that focus on co-location versus co-presence. In describing how naturalness is not merely dependent upon the affordances of the environment, but on what situated actors make of those affordances, what O’Hara and colleagues are saying is, in essence, that the issue of naturalness pertains to social situations rather than to environments.

This argument also applies to the view of social presence as an illusion of non-mediation. In this perspective, actors will only be together when they fail to acknowledge that the meeting is mediated, i.e. when they fail to distinguish the mediated meeting from a face-to-face meeting. However, apart from those cases where the environment is clearly disrupting the meeting, and thereby making it difficult to maintain, it is not obvious how ‘forgetting’ about the medium will necessarily result in an experience of being together. To what extent does it, for example, matter for the sense of being with someone in a chat room if the participants remain aware that they are interacting through a medium? Nowak (2001:4) similarly argues that the importance of the ecological realism a medium can convey has been exaggerated. She writes:

A satisfactory level of co-presence with another mind can be achieved with conscious awareness that the interaction is mediated. People have achieved a sense of another, created friendships, developed communities and conducted business interactions while being very aware that their connections were mediated... Thus, it is likely that mediation will influence the degree of co-presence with another mind, but it would not prohibit this sense.

Here, Nowak is not only arguing that an illusion of non-mediation is irrelevant to a kind of co-located awareness of others, but she also points to its irrelevancy for co-presence (as it is in co-presence we would expect to find the evolving friendships, communities, and business relationships that Nowak mentions).

In addition, if we think of the ‘illusion of non-mediation’ in terms of freedom from environmental disturbances, then we must also take into account that the face-to-face environment is not spared from such disturbances. Actors are often (if not constantly) reminded of the face-to-face environment, for instance when the background music playing at a bar is so loud that one cannot keep a sensible conversation going.

Overall, the kind of ‘realness’ that scholars ascribe to physical environments therefore appears to be exaggerated. It does not, for example, make much of a difference that I am spatiotemporally co-located with a person, if that person is engrossed in a video game or is busy with an instant messaging conversation on the phone. At that moment in time, the video game or the instant messaging conversation is probably more acutely ‘real’ to the other than his and my co-location in the room. Another example of this can be
taken from Puro’s (2002:23) discussion about the use of mobile phones in public spaces.

As someone talks on the phone, one is in her or his own private space. Talking on the mobile phone in the presence of others lends itself to a certain social absence where there is little room for other social contacts. The speaker may be physically present, but his or her mental orientation is towards someone who is unseen (Puro 2002:23).

If this phone conversation took place in the bus queue and I was waiting in line, overhearing what the one person was saying to the other, it is still likely that, in the person on the phone’s social reality I myself only play a peripheral part, while the part of whomever she is distantly engaged with is substantially bigger. Even if we were to assume that certain parts of a person’s current reality could be more or less real than other parts (which is unlikely), spatiotemporal co-location certainly does not prove to be a good measurement. As Berger and Luckmann (1966) would have it, the reality that is ‘closest’ to actors is the zone of everyday life that actors can directly perceive and manipulate, and digitally mediated interactions are in fact directly accessible to the participant in this way. Even if some (expressed) information and immediacy is lost in the process of sending a digital message, and even if the actor cannot perceive the other with her naked senses, she can still, in her ‘here and now’, directly perceive and manipulate the interaction. And so to the person on the phone, the phone call does not take place outside of her ‘here and now’—it is an integral part of it.

On a slightly parenthetical note, there is another danger in the claims about the realness of our lived experience made by the cues-filtered-out proponents. If the realness of environments comes down to how many perceptual channels the actor can perceive her current reality through, would that not also mean that the realities of individuals with impaired abilities to perceive cues are, essentially, always less real or less natural than the realities of people without such impairments? Surely, few would argue that this is the case. A blind person walking down the street is in no less of a real reality than the seeing person she passes, even though her perception of the environment might differ from his.

The Actor as a Passive Consumer of Stimuli

One of the driving questions in social presence theory concerns the relationship between environments and co-location. As I will illustrate in the following, however, the actor himself is overlooked in this model, and the focus remains rather one-sidedly on the environment and its capacity to exert influence. Considering that many social presence scholars describe social presence as a psychological event—such as an emotion or experience—and therefore as something that happens within the actor, this statement might
come as a surprise. However, while I am not arguing that social presence theory is deprived of individuals or human beings in possession of inner lives, I do argue that it is deprived of actors with agency of their own.

The first example of this lack of actors with agency has already been discussed in a previous section. When social presence is described as little more than non-reciprocated or reciprocated awareness of other beings in the environment, it acquires something of an involuntary, and pre-interactional, character. Human beings cannot help but perceive their environment, and the things and beings in it. And to be co-located, a human being cannot help but be perceptible in the environment. Of course, there are plenty of choices an actor can make beyond that—she can choose which environment to enter and thus perceive and be perceived in, and she can, to varying degrees, choose how to appear in the environment and in that way affect how she is perceived. In fact, this last choice has been duly noted in social presence theories that focus on self-presentation.

That the actor has many options as far as co-location goes is particularly true in the case of digitally mediated environments, which are usually accessed deliberately and consciously (and can be chosen not to be accessed at all), and in which the actor has more control over her appearance. Having chosen to enter a chat room, the actor can thereafter decide what user name to ‘appear’ with and to what extent she wants to share information such as gender, age, or ethnicity, about herself. The living physical body, on the other hand, must always be in a physical environment, so the actor’s options come down to what physical environment to be in, and never to whether or not she should be in a physical environment at all. Moreover, the actor has less control over her own appearance in physical environments, as the body will always express embodied cues. Despite that, the actor who is located in a physical environment still has agency and will do her best to take charge of what information her appearance transmits in physical environments (see e.g. Goffman’s 1959 discussion of impression management). Beyond this, an actor can also modify what she perceives by taking action in other ways, in which case she will indubitably influence and alter the way (perceptible) things play out in the environment.

Designers of interactive systems are well aware of the gap between what the designers/researchers think that actors will do, and what the actors actually choose to do. The ingenuity of users thus makes the task of foreseeing how they will use a new technology less straightforward than it may appear, and it is not uncommon that people’s actual use of a technology, and their motivations for using it in that particular way, are vastly different from what the developers had intended. For example, Höök (2006:242) describes how mobile phone users sometimes appropriate their phones’ Bluetooth function for social connection, by giving their phones innovative names that can be picked up by other nearby Bluetooth users.
These names are not pushed to others nearby—they have to turn on their Bluetooth and then search for other Bluetooth devices around—not a very efficient means of communication one might think. And still some users find it intriguing and will occasionally have a look around to see who else is there and how they named their device. This use of Bluetooth is not the one intended by the developers of Bluetooth technology, the only reason you name them is in order to connect, for example, your ear-piece with your mobile phone.

This example goes to show that even when faced with affordances that strongly promote certain actions over others, it does not guarantee that the actor will not opt for more impractical or unexpected choices of action. In addition, the example also illustrates actors’ abilities to appropriate technology to serve some social function, even when the technology can only be pushed to afford a very rudimentary connection to others. In relation to online chat technologies, Latzko-Toth (2010) suggests that the emergence of two distinct interactional chat ‘formats’ is also the result of appropriation. Interaction in online chat rooms (including Internet relay chat), he argues, has emerged as an interactional chat format that is more closely tied to the metaphor of a shared persistent place (i.e. a ‘room’ or a ‘channel’) in which actors readily engage in ‘third place’-like behaviour. The interactional format of instant messaging is by contrast more dyadic and private, and actors ordinarily know each other prior to the interaction. So while the underlying technology is essentially the same, it has been appropriated to afford two different interactional formats: the more site-centred format of online chat rooms and the more self-centred format of instant messaging (2010:369).

A final example of how humans have appropriated communication media to suit their interactive needs is the use of onomatopoetical expressions and emoticons in written messages. These forms of expression are typically used to quickly convey emotion or the intent of a message, which can otherwise be difficult or cumbersome to express in writing, and they can be traced back to long before the advent of the Internet. However, it is safe to say that Internet users have embraced the practice of using nonconventional symbols and expressions to make text-to-text interaction more effective, and this creativity of Internet users essentially represents acts of appropriating the medium to better fit the interactional needs of users.

Nonetheless (and as previously illustrated) the actor is often portrayed as a passive receptor of the external environment in social presence theory, and whatever she might be doing of her own accord goes unmentioned. Indeed, social presence is more often than not described as an environmental affordance or a psychological response or reaction, than it is a form of overt and intentional action in the world. To the extent that the actor himself helps shape the environment or the experience of social presence by actively participating in the proceedings of the situation (or by redirecting her perceptual attention from one object or subject in the environment to another), it re-
mains unexplored in the literature. So the study of co-location—as represented in social presence theory—takes into account the stable affordances of the environment, and the human being’s rather automatic perceptual processing of stimuli in the environment. This is a risky approach, as it overlooks the mutually influential relationship between environments and actors (as described by Gibson 1986), and the relationships between actors. Instead, the approach concentrates mostly on the influence that the environment has upon human beings, and disregards that actors can also exert influence over the environment and others in it.

**Why Togetherness Extends Beyond Co-location: The Example of Chatbots**

As I have suggested, the field of social presence research concerns itself with questions of co-location, which is indeed a necessity for an encounter to come in to being and evolve. Theories of social presence are often detailed, where authors present complex descriptions of how people respond psychologically to cues or information about others presented to them in the environment. The question is, however, whether these theories are really complex enough to fully grasp what it means to be together. As has been described in the literature, the ideal forms, or highest levels, of social presence amount to reciprocated awareness (Biocca & Harms 2002); a media environment capable of transmitting identical social cues as those transmitted in physical environments (Short et al. 1976); or a complete failure of the actor to realize that a mediated encounter is mediated (Lee 2004). But even these ‘ideal’ forms of social presence, I would argue, do not even begin to capture the intricacy of how people come to be together, or what togetherness is, because they do not take co-presence into account.

The example of virtual conversational agents (VCAs, colloquially called chatbots) should help clarify this point. A VCA is a computer system designed for digitally mediated conversation with humans on human terms. But as described in conjunction with the Loebner Prize Contest in the introductory chapter, VCAs famously struggle to make humans feel that they are engaged in human-to-human interaction when they are interacting with a VCA. So what is it that makes the endeavour of passing as human so challenging for a VCA?

To begin with, what makes it difficult for a VCA to pass as a human being in the online chat environment is not co-location. The appearance of a VCA and the means it uses to communicate can be completely indistinguishable from those of a human user. A VCA can be represented by, for instance, a user name, a profile page or an animated avatar, and can use text, graphics,

---

36 To interact on human ‘terms’, conversational agents use so-called ‘natural language processing’. NLP allows a computer program to derive meaning from human (i.e. natural) language input and to also generate ‘naturally’ meaningful outputs in conversations.
speech, animated gestures, or a combination of modalities to communicate with people (in the following I am focusing on VCAs that are designed for text-to-text interaction). The information an actor can draw from just observing other actors in the digital environment therefore does not suffice to determine whether those others are human actors or conversational agents, unless this is intentionally made obvious by the appearance of the other actor (for example, the appendix “bot,” short for robot, is often used in the user names of explicit—or non-deceitful—VCAs to avoid confusion).

In addition to this, a VCA is typically very quick to reciprocate the user. It never has to sleep, eat, clean the apartment, talk to its children, walk the dog or shop for groceries, and can thereby guarantee the user its full attention at all times. It is always online and ready to enter a conversation (except, of course, in case of reboot or technical problems). In addition, a VCA is willing to accept and reciprocate most opening attempts. It will almost always answer\(^{37}\), and it will do it as quickly as is technologically possible. At the level of co-location, therefore, there is nothing stopping a VCA from passing as human.

However, both ‘deceptive’ and ‘explicit’ VCAs do struggle to keep their encounters with humans ‘human’. In the Loebner Prize Contest, where the conversations are time limited\(^{38}\), most computers fail to pass as humans. The contest might, however, not be the best measure of how successful the VCA would be in natural conversations as all (human) participants are here aware of the purposes of the VCA (that is, to deceive), and judges often actively try to catch the VCA out. In natural conversations, it is therefore possible that naïve human users would take longer to discover the true nature of a deceptive VCA than the judges of the Loebner Prize Contest do.

But even in natural settings, users who engage with VCAs that are designed for long-term conversations and relationships quickly get bored of interacting with them (Payr 2010). As Asplund (1987) would have it, the human actor is prone to become asocially non-responsive, simply because she grows tired of the game. This is often explained by the repetitiveness, predictability, and redundancy in the VCAs behaviour (Bickmore et al. 2009). However, as Payr (2010:39) rightly points out, repetitiveness as such cannot be held responsible for the failure to design VCAs for long-term use, as repetition and routine take up a significant part of communication in long-term human relationships. This suggests that while a VCA can successfully imitate repetitiveness, it still fails to be repetitive in the way a human actor would be.

\(^{37}\) Some VCAs can, however, terminate conversations in which the user behaves improperly (e.g. by using offensive language), and can remember that user’s IP-address and refuse to begin a new conversation with any user of the same IP-address.

\(^{38}\) The time limits have varied over the years, from five minutes per conversation to 20+, or 25 minutes.
VCAs also struggle to be coherent, and to understand the context and contextual references. A VCA can give you the correct answer to what 3781295723 times 373581 is in the blink of an eye, but may have trouble telling you how many eyes three children have together, or what the weather is like in the city it just claimed to be living in. Since they do not possess the common sense knowledge humans acquire just by being in the world, and are often unable to read between the lines and decode vague statements (McIntire et al. 2010; Takenobu et al. 2003), their artificiality is easy to detect by those who know what questions to ask.

However, arguing from the standpoint of conversation analysis, Wallis and colleagues (2010) propose that what matters in conversation is not so much that an agent is able to produce logical, sensible and non-offensive utterances, but that, if it fails to do so, it is able to repair the situation. If an agent cannot be held accountable for its utterances, it risks sanctions (e.g. verbal abuse or abandonment), and to be accountable, it needs to be consistent. In this context, consistency means that the agent’s behaviour corresponds with what the other expects, and that way it functions as a form of ‘intuitive’ affordance. As accountability and consistency come down to expectations, this matter belongs to that of the social situation, or rather to how actors understand the social situation. To be consistent, an actor must either have the actual ability to assess what is expected of her in the social situation—i.e. she must be co-present—or she must make it appear as if her actions are accounted for. Not surprisingly, therefore, in many of the examples where chatbots have succeeded in capturing the human’s long-term interest or even deceiving her, it seems as if the chatbot has successfully imitated accountability and thereby co-presence.

For example, when Michael Mauldin (2009:416) created his chatbot Julia to explore and map a text-based TinyMud world and give other users advice and directions, he also decided to give her a personality and a ‘lunar cycle.’ Because of this, Julia was able to avoid questions she did not ‘want’ to answer in an accountable way (i.e. premenstrual syndrome made her grouchy). The following is an extract from one of her conversations in TinyMud (which she entered in 1989), reprinted in Mauldin (2009:416):

Robyn: how are you feeling?
Julia: Terrible.
Robyn: why, julia?
Julia: I have PMS today.
Robyn: Julia, are you a human?
Julia: Is the Pope Polish?
Robyn: yes, Julia, but are you human?
… julia, are you a computer?
Robyn: Julia, are you a woman?
Julia: I am female.
Robyn: and, Julia, are you human?
Julia: Not when I’m having my period, Robyn.
So any personal trait (permanent or temporary) that can explain and excuse rude, inconsistent, illogical, ignorant or repetitive utterances can help a VCA keep the user engaged in the interaction. While Julia exploits the expectations a user may have of how a premenstrual woman behaves\textsuperscript{39}, other traits that can be used to manipulate a user’s expectations of the other’s communicative sophistication are (young) age, language proficiency (e.g. speaking in a second or third language), or an eccentric personality.

Of course, accountability not only comes down to the actor that needs to account for her behaviour, but also to how forgiving the other actor is prepared to be, and what his overall expectations are. Therefore, applying a romantic or sexual frame to the conversation has proven beneficial for VCAs that are designed to pass as humans. This is illustrated in the conversations held between the VCA Jenny18 and one of her returning (naïve) admirers—dom01—on a dalnet IRC channel in 2006\textsuperscript{40}:

\begin{verbatim}
<dom01> waitttttttttt
<jenny18> hm... =)
<dom01> ahhhhhhheyahhhhhhh
<jenny18> mmmMm..
<dom01> fuck
<jenny18> i see
<dom01> i came
<jenny18> hmm
<dom01> yes
<jenny18> ok, good..
<dom01> thanks hey
<jenny18> thank me?? hehh i should be thankin u =)
<dom01> a loved it
\end{verbatim}

de Angeli (2006:2) points out that cybersex is likely to be a winning application for chatbots. That is because it is easier to model conversations that fall within the domain of impulse and passion, as they do not need to follow the normal rules of cognition and logic. In sexual conversations, she argues, deviations from established conversational rules are therefore both expected and tolerated, and the VCA does not have to be very sophisticated. In other words, jenny18 passes as human because she manages to adhere to the rules and expectations of this particular focused social situation.

However, while a human can adapt to different social expectations, and account for her behaviour accordingly, a VCA cannot in the same way read

\textsuperscript{39} Common belief has it that PMS is accompanied by irrationality, moodiness and antisocial behaviour (Chrisler & Caplan 2002).
\textsuperscript{40} Transcripts from Jenny18’s conversations on IRC can be retrieved from the home page of Jake Kaufman (creator of the Jenny18-script) at http://virt.vgmix.com/jenny18/logs/
different situations and frames. In fact, a VCA can only give off the impression that it understands and behaves in accordance with the shared frame, and in that way imitate co-presence. Although it is possible that the conversation between jenny18 and dom01 could still have run its course in the exact same way if jenny18 was human, it is likely that a human jenny18 would have behaved in a different way in an online chat conversation with her parents. By contrast, the VCA jenny18 does not understand that different social situations carry different behavioural expectations; in fact she does not even understand what a social situation is. And so while she may pass as human in situations that are framed as sexual or romantic (because the situation fits her behaviour), her inability to adapt her behaviour to the situation will make it difficult for her to pass (as human) or be accepted in other social situations.

In summary, then, VCAs may not be particularly smart or particularly sophisticated in conversation, but they do—occasionally—manage to ‘lure’ humans into a shared social frame where intelligence, consistency and conversational sophistication are of secondary relevance. Although, in the long run, the interaction between a VCA and a human probably requires additional substance (like the prospect of a long-term romantic affair) to remain interesting to the human, the key to designing a VCA that manages to at all be together with humans is to make it imitate co-presence, rather than co-location.

By this rather long example of VCAs in online chat, and by the previous discussions about the social instinct, the lack of consideration of actors’ own agency, and the focus on physical realism, I hope to have convincingly shown that, if the ultimate goal of the field of social presence research is to understand what it means to be together in a digitally mediated environment, then the focus on co-location in the study of social presence is somewhat misdirected. Co-location, in the sense of being within the perceptual range of cues that signal another person’s presence, is not compromised by the cues offered in an online chat environment, because humans do not need cues of a particular quality or amount to recognize that they are co-located with others. And so the interesting part of our being together in online chat is not co-location as such. Even though we must assume it, co-location in itself does not guarantee that people come to enter a social situation in online chat and it particularly does not guarantee that an encounter will end up being durable and rewarding, as exemplified by the meetings between humans and VCAs. Next, therefore, I will discuss how the microsociological view on co-present togetherness stands up to the challenge of conceptualizing the phenomenon of being together.
Microsociology and Togetherness as Co-Presence

In the second part of Chapter 2, I investigated what ‘being together’ could be said to stand for in theories of face-to-face interaction. Here, a picture of being together as co-presence emerges. According to the literature, co-presence firstly requires that actors are reciprocally aware of, and available to, each other. Passively observing the Queen of England give a speech on television, the arguing couple through their living room window, or a discussion that unfolds in a public online chat room, would not suffice for co-presence, and therefore it would not (in this perspective) suffice for actors to be somewhere together with any of those non-reciprocally observed persons.

At the very least, therefore, co-presence denotes a state of ‘unfocused’ reciprocal awareness that is only loosely framed. In focused co-presence actors must also be overtly responsive to each other and cooperate so as to sustain a single focus of attention. It therefore presumes that actors share a more elaborate frame or understanding of what is going on. To repeat, a definition of the situation, or a frame, represents the meaning an actor ascribes to a particular encounter, and that meaning helps her understand what is going on in the situation and what behavioural norms she, as well as other participants, are expected to follow. While the field would not deny that the existence of environments in which actors can perceive each other is what makes encounters possible in the first place, the emphasis on a shared understanding of the situation suggests that actors need more than ‘just’ an environment to be together in—they need a social situation. So togetherness is here seen as something that occurs within situations—the where of being together is the social situation.

In the following discussion I will be more concerned with focused co-presence (or encounters) than with unfocused co-presence when discussing togetherness in online chat. Assuming that actors are in an encounter, there is still plenty of variation in how drawn into the shared encounter actors are and how consumed they are by it. Like social presence theory—where the interest lies in how we can be co-located with others while we are at the same time physically distanced from one another—the literature on face-to-face encounters is also concerned with different degrees of closeness and distance in its discussions of co-presence. However, the distance that the latter field is concerned with has to do with the distance that can occur between people that are in each other’s close physical proximity, i.e. that are spatiotemporally co-located and thereby not experiencing the distance presupposed in social presence theory.

Spatiotemporal distance—more specifically, a lack of spatiotemporal immediacy in interaction—is usually considered detrimental to all forms of encounters in the microsociological literature. While, for example, Schütz (1932/1967) suggests that the temporal immediacy of a phone call might suffice for actors to enter a we-relationship, the literature generally considers
face-to-face interaction to be the only environment immediate enough to support a social situation in which actors can be together. But even in situations where interaction is immediate, how ‘closely’ actors are co-present may vary. Here, scholars point to various modes of relating, such as involvement and intimacy, as factors affecting the closeness of a meeting.

So how people are co-present, and how closely they are co-present, are shaped not just by how the environment affords actors the ability to perceive each other and transmit social signals, but also by how actors frame, and accordingly decide to behave, in the situation. The social behaviour of actors is shaped by how they understand the situation, which includes how they understand the behavioural norms and expectations that the situation imposes on them, but also to what extent they perceive that other participants share this understanding. In this regard, one important point that Schütz (1932/1967:168) makes is that, when speaking about the closeness in co-presence we are not merely interested in how concrete the other appears to the subject: “[i]t is not only the object […] that is experienced with greater or lesser directness; it is the relationship itself, the being turned towards the object, the relatedness.” In other words, relationship concreteness also involves how the subjects are oriented towards the ongoing relationship and themselves within it. This distinguishes Schütz’ view of togetherness from the favoured ‘object of orientation’ in social presence theory, where the most important object of attention is (with some exceptions) the other person.

The examined literature on focused co-presence is, however, limited in the sense that it only focuses on co-presence in physical environments, and therefore the environment’s impact on situations here translates into the physical environment’s impact on situations. This means that in both of the reviewed discourses (social presence theory and microsociology) a discussion about how digitally mediated environments affect co-presence is missing. In the first discourse (social presence theory) the digitally mediated environment is considered, but only in relation to co-location. In the second discourse (microsociology) co-presence is considered, but only in relation to physical environments.

The cited theories on being together as co-presence are therefore outdated in that they cannot be directly applied to the many modern day social situations that are made possible by the digitally mediated extension of physical environments. Therefore, it is possible that the modes of relating that are described as affecting distance and closeness within face-to-face situations are not applicable or relevant to online chat situations and other digitally mediated meetings. Yet, the interactional emphasis in microsociology is a promising alternative to the perspective on being together as co-location, and might help circumvent some of the identified limitations in the latter approach.

With regard to Goffman’s frame theory, Rettie (2009) makes a similar observation. By adapting Goffman’s idea of a framed encounter (i.e. focused
interaction) to the analysis of mobile phone users’ experiences of phone calls and mobile text messaging (e.g. SMS and Instant Messaging), she concludes that Goffman’s concepts are useful for mediated interactions that are synchronous (e.g. phone and video calls), but less so for asynchronous communications (e.g. e-mail and SMS). Even if they are not spatiotemporally immediate, Rettie argues that synchronous interactions can be understood as focused encounters because of the interactants’ experience of shared time. That is because the determining factor for whether or not actors experience that they share an encounter is that the environment can afford close coordination (through temporal immediacy), while the richness of the information it affords (through spatial immediacy) is less important.

Interestingly, however, Rettie’s findings suggest that synchronicity is not merely a technological affordance, but is also decided by how actors frame what is going on. While text-to-text communication is always less immediate than face-to-face communication, video calls, or phone calls, actors seem to perceive some forms of text-to-text interactions as more synchronous than others. For example, in e-mail correspondence and SMS, actors normally do not expect others to reply instantly, and they do not feel pressured to themselves be constantly attentive to their inbox or to respond as soon as they receive a new message. In other words they expect this form of text-to-text communication to be asynchronous. By contrast, when actors engage in instant messaging conversations, they expect of themselves and others that a continuous focused attention to the interaction will be maintained and that everyone will engage in coordinated interaction for a sustained period. This involves not just an expectation of how quickly actors must respond to messages, but it also concerns an understanding of the encounter as one as having a clear beginning, middle, and end. Instant messaging is thus perceived as a synchronous form of communication that has a certain duration.

As a consequence, the way in which an instant message is composed differs from how an e-mail is composed. In the latter case, the actor will attempt to write down everything she has on her mind before sending the e-mail. In the former case, the messaging follows a more conversational style, with more frequent turn-takings and where actors write shorter messages that do not contain ‘everything at once’. Therefore, the overall content of the instant messaging conversation is less predetermined and develops more gradually throughout the duration of the conversation. For that reason, Rettie suggests that only mediated situations that are perceived as synchronous warrant situational (or frame) analysis, as (actual and/or perceived) synchronicity is needed for actors to share an experience of coherence in the situation. I would like to argue that, in the context of online chat services, it is not only instant messaging that is perceived as synchronous but that this also extends to online chat rooms.

While Rettie’s analysis is primarily focused on mediated interactions that are temporally immediate yet spatially non-immediate (e.g. phone calls), the
following discussion will focus on encounters in online chat environments. Since online chat is technologically asynchronous, yet typically perceived to be synchronous, it makes an interesting basis for discussing digitally mediated togetherness from the perspective of co-presence.

Togetherness in Online Chat from the Perspective of Co-Presence

As described above, co-present encounters happen when actors are reciprocally aware of each other, have an understanding of what is going on, and also perceive that their understanding of the situation is shared with the other. Therefore, for a co-present approach to togetherness in online chat to all make sense, one must begin by assuming that the technology does indeed allow actors to socially interact within a shared frame. According to Rettie, online chat interaction can indeed be considered to be co-present encounters on the grounds that actors typically perceive and treat online chat interaction as focused interaction. Moreover, the technology can be used for a variety of social purposes—to do business, to talk to friends, to find love, and so on—which means that it can be used to relate to others in different ways. This would suggest that actors can also experience greater or lesser closeness to each other in online chat encounters, and that there are more factors than the afforded immediacy that influence this.

While I have argued that social presence theory is too focused on co-location in digitally mediated environments, the following discussion will still strongly emphasize how online chat environments affect co-presence. This might seem counterintuitive, but the rationale is that when environments are examined for the sole purpose of studying co-location, the only thing of interest is the environment’s ability to give off cues about other social beings and how it allows actors to perceive those cues. For co-presence, the environment and what it affords is interesting insofar as it may influence how actors understand and share understandings of social situations in (in this case) online chat, as well as how they negotiate the frame and relate to others within those situations.

In the following, I will therefore discuss focused co-presence in online chat based on how co-presence is understood in the literature on face-to-face interaction. First, I will look at how actors come to understand online chat situations and how they ascertain whether their understandings are shared with other participants. This could be described as a discussion of how co-presence is established and maintained in online chat interaction, or how actors enter, maintain and negotiate a shared frame. Second, I will discuss closeness and distance in online chat togetherness based on the modes of relating described in the microsociological literature. When speaking of aspects of online chat that are general to all types of online chat applications I
will refer to it as online chat or text-to-text interaction. At times I will take a
closer look at specific types of online chat environments—namely online
chat rooms and instant messaging environments—that differ from each other
in some way but still have in common that near-immediate text-to-text inter-
action is the primary mode of communication.

**Framework in the Online Chat Situation**

To briefly recapitulate what was discussed in Chapter 2, one can say that an
understanding of the situation consists of the actors’ understanding of the
situation she is in, what is going on in it, what is expected of her, and what
she can expect of others. It is thus a kind of interpretative frame, through
which everything that is going on in the situation will be given its situation-
specific meaning. However, as pointed out by Goffman (1974), actors are
also ‘doing’ framework in the situation—the frame is something that is man-
aged, negotiated and co-constructed by actors in interaction.

A framework is shaped by whatever information the actor has access to—
hers previous experiences or general knowledge of similar situations, but also
of information that is accessible to her by the environment or by other actors,
before or during the ongoing encounter. An actor typically forms a tentative
understanding of what is going on that is later modified as she accesses more
information. However, to further confirm that an understanding of the situ-
ation is shared with other participants, actors must rely on information gained
through the ongoing interaction. If the behaviour of one person makes no
sense seen through the interpretative frame another actor is using, then the
definition of the situation is probably not shared. Throughout interaction,
actors will look for cues that reassure them that the shared frame continues
to be shared, and, if this information tells them otherwise, they will either try
to modify their frame (i.e. reframe the situation) or break away from it.
Goffman (1974) often notes that, in an encounter, actors are highly motiva-
ted to seek harmony between the behaviour of self and others, and the frame
that surrounds it. That is because behaviour that deviates from the under-
standing of the frame will have negative repercussions for the overall en-
counter.

Since an actor’s understanding of what is going on is linked to what in-
formation she has access to, and since digital communication technologies
alter the actor’s access to information, it would follow that actors come to
understand online chat situations differently than they do face-to-face situa-
tions. To begin with the initial definition of the situation that actors make,
most actors that have some experience of digitally mediated interaction
probably also know that digitally mediated social life is heterogeneous; not
only in terms of how it will allow the actor to communicate (text, images,
video- and phone calls) or how many participants she can expect to meet
(one-to-one, one-to-many, or many-to-many), but also in terms of the nature
and function of different social applications and websites. Online dating
sites, social networking sites, online chat rooms, discussion forums, virtual worlds, instant messaging applications, and online collaboration platforms all serve different purposes, allow for different forms of interaction, and impose different behavioural norms. In Goffman’s terms, the digitally mediated environment does not just hold one single overarching social frame, but many different ones.

It is probable that an actor does not enter an online collaboration platform for the same reasons she enters a dating website or an instant messaging environment she shares with friends. She will expect different things, and likewise feel that the expectations of her behaviour differ, from and between each of these ‘places’. Instant messaging environments are, for example, primarily used by people who have already established some sort of relationship with each other. For a user to text another user on WhatsApp, she must already have that person’s phone number saved in her phone’s address book, and to contact someone on Google Hangouts the user similarly must have that someone added to her “circles”. To engage with others in the online chat room Zobe, by contrast, users do not need to have any prior knowledge of each other; in fact it is described as a service that allows users to meet new people under anonymous conditions.

In Zobe and other online chat room services (like ICQ chat) there can be a selection of ‘rooms’ for users to enter. Such rooms are often organized based on interests, age-span, or nationality, so that a user who is interested in talking about, say, parenting can enter the Parenting Room and expect to find others with similar interests there, while a Swedish user can enter the Swedish room and expect the ‘official’ language in this room to be Swedish. In that sense, the organization of online chat is similar to how face-to-face interaction is organized, where different place-bound situations are inscribed with different meanings, and in which certain frames are more readily applied. Instant messaging is more private, while online chat rooms are more public. In one online chat room one might find people who are interested in speaking about sex and romance, and in another one finds conversations about children and parenting methods. The rules for what is appropriate in private versus public online chat, or in a chat room focusing on dating versus parenting will of course vary, just as the expectations of users will vary depending on the nature of the online chat service. The same principles extend to other social arenas on the Internet, which means that, as Zhao and Elesh (2008) observe, digitally mediated social life is highly regionalized.

The frame provides the basis for some of the actor’s expectations, but the situation also consists of particular others of which the actor has particular expectations. These expectations are formed by what information the other expresses (or gives off) and communicates (gives), and actors typically regard the information given off as more representative of who the other ‘truly’ is (Goffman 1959). Environments play a part in how people form their expectations of others in that they dictate what information actors can pre-
sent themselves, and the level of control actors have over the information they communicate and express.

In particular, the information that can be gained from an actor’s mere appearance differs between online chat situations and face-to-face ones, because the online chat environment does not afford embodied appearances (at least not in the physically embodied sense). However, different online chat environments allow or require the user to volunteer different types or amounts of personal information, and services can be more or less strict about verifying whether users are submitting the correct information. Several online chat rooms, such as Zobe and ICQ chat, are designed so as to allow their users to remain relatively anonymous throughout their visit. Zobe, for example, only asks the user to select a user name and to provide information about her age (she can choose between 13-17, 18-23, and 24+) and gender (male or female only) but it does not require the user to further verify this information. By contrast, other services that allow for instant messaging between strangers will ask the user to in some way certify that the information she volunteers is correct.

The popular dating site Tinder for example prompts the user to log in using her Facebook account, so that her Tinder profile is based on the personal information she provides on Facebook. In that way, the application supposedly ensures that users are who they say they are, which can otherwise be a nuisance in the context of online dating. To take another example, the WhatsApp user must provide her correct phone number in order to access the service, but she can then choose a profile picture and a ‘status’ message that will be shown to all her contacts. She can also choose to not upload a profile photo, and she does not have to provide a status message other than the default one (i.e. “Hey there! I am using WhatsApp”). Unlike many other applications, however, WhatsApp users cannot control what name they will go under in the WhatsApp environment, as the service draws that information from the user’s contact phone books. So if Miss A’s friend has stored her number in his contact book under the name of Mrs. B., Miss A. will also appear as Mrs. B. in the friend’s WhatsApp application. The way in which actors appear in online chat environments thereby varies across applications, but they all have in common that the appearance must be crafted by information provided through text, and in some cases also through photos and images.

Just by observing another person in a face-to-face environment, an actor will draw many conclusions about what to expect of that person. Those expectations are based on what information this person’s body is continuously expressing (about, for example, age, gender, socio-economic status, ethnicity, disabilities, and so on). While actors can to some extent alter their physical appearance to change how others see them, they can only do so within certain limits. An adult will, for example, find it difficult to pass as a child,
and a person who only has one leg will struggle to pass as a person without visible disabilities.

Of course, in online chat encounters, actors also form expectations about each other upon the ‘first look’ at each other’s digital self-representations (e.g. the already mentioned user names, profile information, and profile photos), but the difference is that users have more conscious control over what their digital self-representation communicates, for instance with regard to the various categories to which they belonging (such as gender, age, nationality, and so on). In addition, the actor also has greater power to decide whether she wants her appearance to reveal a lot or a little, while her physical appearance can never be altered to reveal more or less information—it can only be altered to reveal different information. For example, an actor that appears in the online chat room solely by a user name gives less information than an actor that appears with a user name, a photograph, and some additional information (e.g. gender and age). But an actor cannot choose to appear in a face-to-face environment only by her name (or only by her name and age, for that matter)—she must enter it with her full body, which will constantly express to others ‘who she is’ even when if she is not consciously trying to communicate this information.

So far, I have been mostly concerned with the tentative expectations an actor forms of a person she has no prior experience of interacting with. But the differences in how environments shape the appearance of actors are also influential in encounters between people with a relationship history. That is because the appearance of the physical body can also express plenty of information about the actor’s mood, health, and state of mind, and this information will be used to form expectations about the situation at hand. An appearance in the online chat environment will not express any indication of an actor’s current state of mind, emotions, or health unless the actor actively communicates this, and by the time that communicated message reaches the other, the actor’s feelings might have changed. Some applications—like WhatsApp—will allow the actor to change the mood or ‘status’ of her appearance (in this case by writing something in her status line) that can communicate to others that the actor is feeling happy, content, sad, or angry. However, in order for such an appearance-attachment to accurately reflect the current mental and emotional status of the actor, the actor would have to consciously and frequently update it, something that is done automatically by the physical body in face-to-face situations. If the issue in face-to-face situations can be that the actor feels her appearance reveals information that she does not want the other to have access to, the problem in online chat situations may be that the actor feels her own or others’ appearances do not reveal enough information.

In this chapter I have made a distinction between unfocused co-presence as something that precedes overt interaction (but also continues throughout it), and focused co-presence as an interactional event. The first impressions
and expectations I have discussed so far are depicted as ‘pre-interactional’ in that they can be formed before actors actually enter into an encounter. However, I believe this discussion is still important, as actors will use first impressions and expectations to define the situation they are about to enter. Moreover, for the co-location described in social presence theory, the actor’s experience of others remains focused on how the actor gets the impression that someone else is there with them.

By contrast, the experiences discussed here (which also go in line with the view of social presence as self-presentation) concern how the actor forms impressions of who is there, what state this person is in, and what can be expected of him or her in subsequent interaction. In the previous discussion of co-location in online chat, I also mentioned that a lack of non-verbal expressions makes it hard for actors to establish that there exists a mutual availability and willingness to interact before they take the next step and engage in a focused encounter. So, while the actor can form an idea of who is there prior to the initiation of text-to-text interaction, she can only assess whether whoever is there is also socially available by making overt attempts to interact with this person.

In order to open an online chat encounter actors must therefore skip the phase in which the risk of later being overtly rejected by the other is managed and minimized. In face-to-face interaction, such ‘risk-reduction’ is normally done through actors subtly and reciprocally signalling that they are willing to continue into interaction. Instead, actors must use more straightforward invitations in online chat, to which a rejection (e.g. silence or a clearly stated ‘not interested’) may be more obvious and potentially more embarrassing, particularly if there is concrete evidence that the other has actually opened and read the message. For example, in WhatsApp and Google Hangouts the user will be made aware of whether the receiver has read a sent message or not (and at what time she read it), and so it may be harder for the user to dismiss silence on the part of the other as a failure to receive or open the message.

Before proceeding to discuss what happens with actors’ understandings of the situation after they have begun to openly cooperate in text-to-text interaction, I wish to discuss how the affordances of the online chat medium affect the actor’s overall understanding of risks and uncertainties in the situation. The already mentioned difficulty in covertly confirming whether or not the other is interested in interaction is merely one example of how the actor accounts for uncertainty differently in the context of online chat than in other interactional contexts. In addition, the first impressions and expectations an actor have of others in online chat environments are made up of quite different information compared to face-to-face environments. This is because, as already discussed, the physical body is not making an appearance in online chat environments, except in its represented form. The representation can be
carefully composed by the user, where she can simply leave out information she does not want others to base their impression of her on.

By contrast, in face-to-face interaction, an actor must go to much greater trouble if she wants to conceal, hide or even alter some information that her body is, by default, continuously and immediately expressing. For example, information about one’s gender can be easy to leave out in an online chat representation of oneself (except when providing this information is prompted), while most actors would have to go through to a great deal of trouble to have their physical bodies pass as genderless. In addition, the comparatively low immediacy of online chat environments and the way in which a self-presentation may linger in it, means that the information an actor is gaining about another person may be out of date. A profile picture on WhatsApp could have been taken ten years ago, and if the user does not actively replace it with a newer photo, it can continue to represent her for another ten years to come. By contrast, the physical bodies encountered in shared physical environments are always presenting themselves in their current states.

While the first impressions we gain of others in face-to-face environments can, despite the immediacy and richness afforded, still result in an incorrect understanding of who that person is, what state she is in, and what to expect of her, the absence of the body in online chat environments further complicates this understanding. In effect, the actor’s perception of uncertainty with regard to what to expect of other participants should be increased in the online chat situation. According to Goffman (1959), the more control the other is perceived to have over his appearance (i.e. the less information he gives off as opposed to gives), the more uncertain the actor might feel that his understanding of the other is correct. As online chat environments effectively cut off many of the expressions that the actor would normally use to determine whether what the other is purposefully conveying is valid, actors must deal here with other uncertainties and different risks compared to those of face-to-face environments.

But as the impression control an environment affords cuts both ways, the heightened control that one actor has also extends to other actors. So even though the environment may grant actors less access to involuntarily expressed information about others, and thereby less control over who the other is, the state he is in, and of how he has received information, the actors simultaneously have more control over what information they themselves send out. In this particular sense, actors have less control in face-to-face environments than in environments deprived of physical bodies, because their physical bodies are constantly giving off expressed information. Information about oneself will therefore be revealed whether it is intended or not, giving the other a better chance at acquiring knowledge that is not explicitly verbalized (i.e. communicated). Of course, in all forms of interaction one might send off unintended information to the other, as communicated information always also expresses information, but the fact remains that manag-
ing what information one sends off becomes easier when actors cannot perceive each other’s physical bodies.

So while the actor may have a sense of certainty that she is being perceived exactly the way she wants, she would simultaneously feel uncertain that the information she receives about the other is an accurate reflection of who that person really is and of what can truly be expected of him, as well as uncertainty about whether his responses are genuine. Online chat environments allow the actor heightened control over her own appearance and expressions, yet less control over others’, and therefore call for new strategies of uncertainty reduction. They also potentially call for shared frames that encourage actors to be less self-censoring.

As the educational perspective on social presence illustrates, a problem in the online classroom is that students do not volunteer enough information (in this perspective, they struggle to successfully convey a social presence or a self-presentation). Therefore, the increased impression control of an actor is not only a problem for the onlooker, but may equally well be problematic from the perspective of the self-censoring actor (in this context, his learning progress is believed to decrease).

Applications can also be designed to reduce uncertainty from the outset, which was earlier illustrated by the example of how, on Tinder, actors are forced to build their appearance on Facebook data. Since on Facebook, a user’s contact list typically contains friends, acquaintances and even family members that the user has met face-to-face, this supposedly creates a social barrier against providing personal information that strays too far from the truth or against users publishing photos of someone else and then claiming these photos represent them. Of course, there is always the possibility that a user may create a Facebook profile that is entirely fictional, and then use that to access Tinder; however, the overall belief is that this function decreases the number of occasions on which users access Tinder with completely made up self-presentations.

Moving on, there is an additional part to understanding the situation—namely the understanding that one’s understanding of the situation is shared with others. To arrive at such an understanding, the actor must to some extent be able to take the perspective of the other, meaning she can imagine how the other actor understands the situation. The ability to take the perspective of the other is, of course, something that an actor learns very early on in her socialization (Mead 1934, Cooley 1902), and the question is rather how the online chat environment influences how she comes to understand that the other shares her understanding of the situation.

Three ways in which online chat environments influence how an actor takes the perspective of the other are: how the environment allows the actor to 1) know who the other is (e.g. categorical information, but also more personal information), 2) know how the other reacts to the actor, and 3) know the point of view of the other. Categorical understandings of others have
been discussed above, and can perhaps serve to make the actor feel her understanding of the situation is likely (or not likely) to be shared with the other at the outset. However, the adequacy of this feeling is only ever tested in interaction, and here access to information on how the other reacts to the actor and the situation in general is what matters. From the way the other behaves the actor typically monitors whether the shared frame is, and remains, shared. This is done, for example, by assessing the way the other reacts or responds to what the actor is doing, or the extent to which the behaviour of the other deviates (or does not deviate) from what the actor understands as the behavioural norms of the situation.

Immediacy and the amount of expressed information afforded here make a difference, because they give actors access to each other’s spontaneous reactions. Since they provide the actor with continuous input from the other throughout interaction, spontaneous reactions are typically used to confirm whether the frame is still shared or whether it is threatened in face-to-face situations. By contrast, in the more sequenced interaction of text-to-text situations, the actor first types a message and sends it. Then the other reads and reacts to the message, types her response (which might not correspond with her original reaction), presses ‘send’, and only thereafter does the actor receive a first response from the other. In other words, the actor does not get continuous feedback, and when she finally receives feedback she cannot be entirely sure it corresponds with the spontaneous reactions of the other. While the actor might experience that she has more freedom to formulate a response that truly corresponds with what it is she wants to communicate, the problem here is that she cannot as quickly assess whether her message is being received the way she intended it to be. But to substitute for the loss of immediately expressed information in text-to-text situations, actors have adopted new habits to ‘express’ themselves, for example by the use of smileys, onomatopoetical expressions, and different writing styles. These appropriations of the written language serve to ensure that the other receives information in the way and with the tone it was intended, and in this sense they serve a function similar to that of the information expressed by the physical body (through for example tone of voice, facial expressions, and body language) in face-to-face interaction.

The way in which the immediacy of information flow affects an actor’s understanding of the situation is thereby linked to the level of control participants have over what information they give off to others and how they give it off. One reason for this is that interaction of low immediacy, which typically takes the form of written communication, leaves a lingering mark that can be observed by the actor in much the same way as the other will observe it. Another reason would be that the time delay characteristic of some low immediacy environments (such as online chat environments) makes it possible to correct and perfect information before it is put on display for the other to interpret.
In situations of high immediacy, by contrast, there is no way an actor can take an outsider’s view of herself. Even though we could probably use tools such as video streaming to watch ourselves in interaction, observing ourselves observing the other observing us can only happen retrospectively. Moreover, we cannot split our attention between self-attention and other-attention and at the same time maintain a high-immediacy encounter, as high-immediacy other-orientation is characterized by synchronous reciprocity of attention. That is, my observing you observing me observing you observing me happens at one and the same time, and there is no way I can take a break and observe myself observing you without reducing the immediacy of the encounter. Therefore, it is harder to control and monitor the information we communicate and express in high-immediacy encounters. This perception of one’s own and others’ control over information in the environment forms an important part of the understanding of what is going on, not least because it, as previously mentioned, relates to the perception of risks—in particular the risk of losing face (Goffman 1959)—in the situation.

A last point to be made is that being in an online chat environment does not mean the actor at the same time exits the physical environment she is in. Rather, the online chat environment opens up a new space in the physical environment that can be shared with people that are simultaneously located in other physical environments. To take the perspective of the other in text-to-text interaction therefore involves understanding, and perhaps accepting, that the part of the other’s perspective that involves her physically defined situation is out of the actor’s reach. While the other can of course both describe her physical location and show images of it, and so give the actor an idea of what she is experiencing, there is still a significant difference between the idea an actor can have of what is currently within the perceptual field of the other in text-to-text interaction, and the idea she can have of what the person she is interacting with face-to-face is perceiving. That is because she only shares full spatial immediacy with the other in the latter case.

There are many implications of this. If the other, for example, suddenly goes quiet in text-to-text interaction, the actor typically has very little information to rely on in order to make sense of the silence. Perhaps the other just witnessed a car crashing into another car, perhaps she is struggling to come up with words to express that that last message made her intensely happy, perhaps she went to the bathroom, perhaps she has lost interest in the conversation, perhaps her mother just called her on her mobile phone, or perhaps she is quiet because she took offence at something the actor just said. In face-to-face interaction, silence on the part of the other can of course still be difficult to understand, but the actor indubitably has more information to work with and thus stands a greater chance of narrowing down what appears to be the most likely reasons for it. What this example of silence also shows is that digitally mediated environments, particularly those low in spatial and
temporal immediacy (like the online chat ones), make it impossible for actors to truly share an undivided point of view.

In summary then, the understanding of the situation, or the frame, is highly related to how the environment allows information to flow in the situation. The most notable ways in which the online chat environment differs from the physically mediated environment, is that it limits the extent to which the expressions of an actor’s physical body are made visible to others, and the immediacy with which communications (and resulting expressions) are transmitted and received. I have here discussed how this affects the actor’s understanding of the situation and the other in the situation. A question that remains to be answered is how the conditions for the activity of framework in online chat situations affect togetherness. Do they, like social presence theory would have it, make it harder for actors to be together or get close to each other? Or can actors work around the fact that the framework in online chat is not the same as the framework in face-to-face interaction, and in so doing successfully be together anyway? Some answers to these questions will be found in the following sections, where I discuss how involvement, mental closeness, and intimacy can be understood in relation to online chat environments.

Getting Closer in Online Chat Situations
In the reviewed literature on face-to-face interaction, it is suggested that it is not only the spatiotemporal immediacy of the interaction that serves to bring actors closer together or further apart; so does the way in which actors relate to each other. Three modes of relating emerged in the literature, of which one has to do with actors’ involvement in the situation, the second with the ‘mental closeness’ an actor experiences towards another actor, while the third concerns the intimacy with which actors relate to each other. In the following I discuss how these modes of relating combine with the affordances of online chat environments. Of particular interest is whether involvement, mental closeness, and intimacy can serve to reduce some of the social distance created by the relatively low immediacy in text-to-text interaction.

Involvement
Involvement is described as a mode of relating that concerns how actors allocate and sustain attention to an ongoing situation. Shared involvement is (to some degree) a requirement for a focused encounter to remain focused, as actors must agree to maintain their attention to the situation. In its spontaneous and most intense form, involvement is depicted as complete absorption in the situation, where the actor becomes one with the interactional flow and—if only for a short while—ceases to consciously reflect on what is going on. The actor is thereby entirely focusing on the ongoing interaction, which would not include paying attention to internal thoughts or reflecting on what is happening in the situation.
The characteristics of text-to-text interaction would suggest that spontaneous involvement in the situation is difficult to achieve. In many online chat environments, the participation of actors must be consciously communicated, and expressions only appear as a side effect of those communications. In addition, the comparatively low immediacy of responses in these environments invites reflection because of the short delay between what one actor says and the next actor’s reply. Of course, an actor can potentially be spontaneously involved in a text-to-text interaction if he types and sends messages as quickly as he possibly can, and in so doing continuously writes what is on the tip of his tongue (so to speak). If the other participant does the same, then it is perhaps possible to talk of a type of joint and non-reflected involvement in the situation. But it is hard to see how, in a scenario like this, actors can interweave their spontaneous contributions with those of the other, so that each contribution meaningfully connects to what has previously been said. In fact, spontaneous involvement in an online chat situation might very well be perceived as spamming, and thus as an inappropriate disturbance. It would seem that while joint spontaneous involvement can here be represented by a flow of letters and words, it cannot be represented by a conversational flow. Of course, each contribution is still meaningful in the sense that it signals that there is another social being in reach, but if these contributions do not come across as reciprocal responses to the contributions of other participants, then it is uncertain if we can describe this situation as a focused one.

If, in face-to-face interaction, actors can read each other at the same time as they are themselves expressing or communicating information, this is arguably more difficult in text-to-text interaction. When the actor’s attention is devoted to producing a continuous stream of text messages, there is not much room left for interpreting what the other person is saying. To the extent one can speak of joint spontaneous involvement in text-to-text interaction, then, it appears as if this involvement is mostly based on quick exchange of messages, while the contents of those messages are less important. This resembles how Simmel describes sociable association, where what actors are saying or contributing to the encounter matters less than that they are participating in a free and playful manner. But while it appears achievable to be spontaneously involved in sociability in text-to-text interaction, it is harder to imagine how joint spontaneous involvement could occur when the situation is framed differently, for example as a business meeting.

But Goffman also describes involvements that are less all-engrossing and where the actor can, alongside his main involvement, maintain side involvements. Here, the influence of the environment would concern how much information is available for the actor to be attentive to, but also how much information is available to assess whether others are maintaining their involvement obligations. As we have seen, digitally mediated environments are often depicted as environments that limit the range of perception (for
example by providing less ‘rich’ social stimuli). However, when it comes to the total amount of things an actor can be attentive to, whatever an online chat environment offers is added to the things already available in the actor’s physical environment (as the former environment is embedded in the latter). Actors are bound to physical environments by their bodies—which, by the way, are what make it possible for an actor to be attentive—and so ‘going’ to an online chat environment does not mark a simultaneous exit from the physical environment, the way travelling between different physical places does. While it may be true that the social cues that can be transmitted in online chat are less rich than those that can be transmitted in face-to-face interaction, it is still true that no matter what digitally mediated environment an actor accesses, the total amount of possible stimuli she has access to will increase when she does so. So if two people are meeting in an online chat situation, the informational situation they share and can be mutually attentive to only constitutes a small part of the entire informational situation that each of them are in.

When an actor goes online, then, the stimuli she can be attentive to comprise both digitally mediated stimuli and physically mediated stimuli. If she happens to be co-located with another person, this other often only shares one part of the totality of her informational environment—the physical part or the digitally mediated part. For example, a couple can be physically co-located in their home, at the same time as both are engaged in text-to-text interactions with other actors. So they are physically co-located, but not digitally co-located, with each other, at the same time as they are digitally co-located, but not physically co-located, with the people they are interacting with digitally. Of course, actors can also both be physically and digitally co-located (for example by being jointly engaged in a conversation on Zobe, while at the same time sitting next to each other on the living room couch). But when actors are digitally, but not physically, co-located, it is the case that even though (compared to a pure face-to-face environment) the online chat environment offers more stimuli to which each actor can be attentive, it still offers less information on how the other person is allocating her attention. This is due to the limited spatiotemporal immediacy and point of view afforded by online chat environments, and was discussed previously.

However, if, in face-to-face environments, an involved actor expects other participants to maintain their interaction involvements as dominating main involvements, the same level of involvement is often not a reasonable expectation in text-to-text interaction, both because of the time delay between conversational turns but also because it is usually not practically possible to monitor whether the other is in fact directing the majority of his attention to the situation. Though interaction through a medium such as instant messaging can get intense, in the sense that both participants are fully attentive to the conversation and reply as soon as they can, it is perhaps more often the case that the interactional flow is frequently interrupted and less synchro-
nous than it can be. After sending an instant message, the actor will probably not sit passively in front of her screen waiting for the friend to reply; it is more likely that she will use the waiting time to visit other web pages, check her e-mail, make a coffee, send an instant message to someone else, and so on. However, it is of course also possible that the actor will in fact give her undivided attention to the first conversation, in the sense that she uses the time between replies to do nothing but focus on the conversation as it is printed on the screen. Yet it is indubitably true that the time delay here invites her to use the spare time on other involvements.

In line with this, Rettie (2009) maintains that text-to-text interactions such as instant messaging do come with involvement rules and obligations. But while the expectation on actors to maintain sustained and prolonged involvement in the situation is similar to face-to-face situations, it is generally not required of participants in online chat situations to give their full and undivided attention to the ongoing conversation at all times. Instead, the way in which these situations are framed allows the actor to maintain several involvements at once, even if those other involvements are also rather engrossing.

Online chat environments thus facilitate participation in several interactions at once. For example, an actor may simultaneously be keeping up two separate conversations in different chat rooms, ‘tabbing’ from one conversation to the other. After a while her mother may call her on Skype, so that she is to some extent involved not only in two, but three different conversations at once. As Turkle (2012) worrily notes, this behaviour of participating in several distinct interactions at once is not limited to digital interactions only:

At home, families sit together, texting and reading e-mail. At work executives text during board meetings. We text (and shop and go on Facebook) during classes and when we’re on dates. My students tell me about an important new skill: it involves maintaining eye contact with someone while you text someone else; it’s hard, but it can be done.

Here, what the environment affords in terms of immediacy makes all the difference. In text-to-text conversations, responses can never be as immediate as face-to-face interaction is, and thereby the expectations of how quickly the other should respond are lowered. In all cases of maintaining several involvements at once, it is unlikely that the actor is fully attentive to any of the conversations (as research suggests that the human brain cannot really multitask41), but she may be able to be sufficiently attentive to each of them

---

41 See Ophir et al. 2009 who compared how ‘heavy’ and ‘light’ media multitaskers performed on various cognitive tasks. They found that heavy multitaskers’ performance was worst in all areas, including the one that tested participants “task-switching ability” (i.e. multitasking-skill).
for her involvement to pass as appropriate. Considering how digitally mediated interaction is now integrated in our daily lives, it is possible that our understanding of what is appropriate allocation of attention in face-to-face interaction has also undergone changes, where we are more tolerant of actors splitting their involvement between us and some other person on WhatsApp or Zobe.

However, Rettie’s (2009) interviewees still describe how talking to someone on the phone while being physically co-present with another person is inappropriate, while sending instant messages could pass as acceptable behaviour in the face-to-face situation. The distinction made here, Rettie suggests, has to do with how much sustained attention the actor must allocate to each task. The phone call requires the actor to maintain some level of attention to the conversation for the entire duration of the call, meaning that her attention to the physically co-present other will be compromised from the beginning to the end of this imposing interaction. By contrast, the text conversation puts expectations on the actor that are more congruent with the expectations imposed on her by the face-to-face situation, as text messages typically do not require that participants maintain undivided attention to the conversation. Therefore, the actor can here successfully manage both her obligations in the face-to-face situation and her obligations in the online chat situation, and in so doing can please both the face-to-face other and the digitally mediated other.

However, more ‘immersive’ digital environments like the MMOW Second Life, or MMORPGs like World of Warcraft make higher demands on the users full, and undivided, attention. As they are designed to mimic spatial environments, they also mimic the consequences involved with moving from one spatial environment to another—namely that one will miss out on what is going on in an environment one has turned one’s back on. And so virtual worlds are not as forgiving when it comes to dividing one’s attention between different involvements, or when it comes to being absent from the environment for some time. For example, interaction in a virtual world is near-immediate, meaning an unannounced toilet break in ‘real life’ will be noticed and will disrupt what is going on. Moreover, the social life of the virtual world does not pause just because the user logs off, and the user cannot directly access what has happened during her absence at a later point. The same is true for many online chat rooms that do not store conversations, and where the user can only access information that is written while she is logged in to the environment.

So logging out of an online chat room or Second Life is thus more akin to leaving a vibrant party early than it is to logging out of Facebook or Google Hangouts. In the first two cases, life goes on without you when you leave and you can only get an idea of what happened after you have left through second-order representations of it. In the case of Facebook or Google Hangouts, life also goes on without you after you have logged off (i.e. peo-
ple can still post information on their Facebook walls, or send you instant messages), but the next time you log in this information will present itself to you in pretty much the same way as it would if you had never logged off. In order not to miss out on something, Facebook and Google Hangouts do not demand your undivided attention the way online chat rooms, virtual worlds, and face-to-face environments do.

So when comparing the demands of immersion in online chat rooms and virtual worlds with other digitally mediated environments (such as social networking sites or instant messaging services), it becomes clear that the temporal structure of the environment is bound to affect participants’ allocation of attention, and thereby also their involvement in the situation. That a ‘weakened’ temporal structure affects involvement and involvement obligations has also been pointed out by West (2000) who suggests it will become harder and harder for people to distinguish between main involvements and side involvements; and between dominating and subordinate ones. Given the discussion above, this observation rings true. Seen as an all-dominating main involvement, ‘joint spontaneous involvement’ is moreover what seems to take the greatest hit in environments where the temporal structure is weakened, as it is in online chat environments.

Goffman (1963) states that involvement in a situation is subject to regulations, which means there are behavioural rules for how co-present actors should be involved. Too much, and the actor disturbs other participants’ involvement; too little, and the actor comes across as indifferent and distanced. In line with previous findings (e.g. Rettie 2009), the discussion so far suggests that involvement obligations of online chat situations would by necessity diverge from those of face-to-face situations, because actors cannot realistically be involved in online chat situations the way they can be in face-to-face situations. For example, even if an actor is completely absorbed by a text-to-text situation, she cannot continuously communicate her full involvement in it without risking ruining her own (because when she is focused on typing replies she cannot focus on other things going on in the situation) and perhaps also others’ involvement (because excessive texting, so-called spamming, may be perceived as deviant behaviour).

Compared to face-to-face environments, it should also be more difficult to judge how involved others are in the online chat situation. The actor can potentially look to how well the other is staying on topic and how regularly she contributes to the interaction. But in order to confirm that the other is still sharing her involvement in a face-to-face situation, she does not need the other to be on topic or even consciously communicate anything. In face-to-face interaction, and possibly also in video- and phone conversations, one can be silently involved, while extended silence is (if anything) a mark of misinvolvement in text-to-text interaction.

Recalling Goffman’s list of misinolved behaviours (i.e. external preoccupation, self-consciousness, other-consciousness, interaction consciousness,
autoinvolvement and over-involvement), only a few of these would translate to the online chat situation. External preoccupations (e.g. watching television while being committed to an instant messaging conversation), self-consciousness (e.g. reflecting upon whether the last message sent will impress the other negatively), other-consciousness (e.g. being overly aware of how many typos the other person is making) and autoinvolvement (e.g. picking one’s teeth) are all behaviours that do not necessarily have to influence one’s own or others’ participation in a disembodied conversation, as they do not need to leave visible marks on what is written or otherwise communicated. That is, the effects of these misinvolvements may never materialize in the shared online chat environment, through communications that are e.g. vague, ambivalent, inconsistent, or unengaged, and thereby they do not have the power to disrupt the situation in itself.

However, interaction consciousness and over-involvement may potentially disturb the overall involvement in online chat situations. As was mentioned earlier, joint spontaneous involvement would, in the online chat context, probably take the shape of over-involvement, where the spontaneously involved actor will struggle to make coherent and logical contributions to the shared conversation. In addition, it is possible that an actor can get overly worried about how the text-to-text interaction is going (i.e. get interaction-conscious), and thereby fail to appropriately participate in the situation. In instant messaging, an actor may for instance spend so much time searching for the right words to communicate exactly what she wants to say that, by the time she finally sends the message, the other person may have already changed the topic. This would then potentially break the natural flow of the conversation.

As it is so tied to spatiotemporal immediacy, involvement is perhaps not an optimal measurement of how closely or distantly actors are together in online chat situations. It is clear that there are many differences between involvement in ‘pure’ face-to-face situations compared to online chat situations, and the degree of involvement in face-to-face encounters appears to be more strongly related to the degree of closeness between actors than it is in online chat encounters. Joint spontaneous involvement is, for example, is probably not a mode of relating that serves to bring actors closer together in the text-to-text situation, as, in practice, an actor who expresses spontaneous involvement has been linked to disrupted perspective taking, goal formulation and other “metacognitive” activities (Cegala et al. 1982). Cegala and colleagues (1982) further argue that an individual experiencing low involvement will appear distracted, preoccupied with other thoughts or goals, unmotivated to interact, unresponsive to others, uncertain, and withdrawn from immediate social context. In speech, the individual comes across as vague, ambivalent, inconsistent, unengaged, and prone to misunderstanding. We would also expect that improper delays (where what is deemed “improper” would depend on the situation and the expectations it raises) and an inability adjust to the other would signal that the other is not fully involved in the situation.

---

42 Low involvement has been linked to disrupted perspective taking, goal formulation and other “metacognitive” activities (Cegala et al. 1982). Cegala and colleagues (1982) further argue that an individual experiencing low involvement will appear distracted, preoccupied with other thoughts or goals, unmotivated to interact, unresponsive to others, uncertain, and withdrawn from immediate social context. In speech, the individual comes across as vague, ambivalent, inconsistent, unengaged, and prone to misunderstanding. We would also expect that improper delays (where what is deemed “improper” would depend on the situation and the expectations it raises) and an inability adjust to the other would signal that the other is not fully involved in the situation.
involvement in a text-to-text situation would probably not be able to maintain a sensible conversation. The lowered immediacy, the increased amount of information to be attentive to, and the lowered control over other actors’ involvements would suggest that, in online chat environments, it is easier for actors to fall into what in face-to-face encounters would be considered misinvolvement. However, there is reason to believe that actors adapt their expectations of their own and others’ involvements to suit the conditions of online chat encounters.

Since the other’s involvement can only be measured by how he overtly and continuously expresses what he is attentive to in each moment, the question remains of to what extent involvement makes an actor feel closer to someone in online chat encounters when there is so little evidence to support that her own involvement is matched by the other’s. But perhaps involvement can also, in text-to-text interaction, be expressed through what is communicated: for example through how well a response follows another, through the questions asked, or through communications that show that one has gone the extra mile to author a message. While such communications cannot tell the actor what the author of the message is currently involved in, they can tell him something about how involved and invested the author was in the communication when he was writing it.

This has some support in the hyperpersonal theory proposed by Walther (2007:2543). Walther argues that digitally mediated interactions—at least the asynchronous types, such as text-to-text interaction—can actually afford greater intimacy and allow actors to establish close relationships much quicker than they can in face-to-face interaction. This is because these environments afford actors the opportunity to a) spend “almost unlimited” time to edit and refine their messages before sending, b) conceal involuntary cues (expressions) so that the actor can present his ideal self, and c) spend less attention on environmental scanning and non-verbal management. Instead the actor can concentrate fully on perfecting his messages. These factors allow the actor to be mindfully involved in the conversation and in the composing of a message, which in effect (it is hypothesized) makes actors feel closer in the digitally mediated situation.

That the time and effort that goes into the composing of a message could matter to the closeness between actors is an idea that bears a resemblance to Mjöberg’s (2011) discussion of ‘mental closeness’, which will be discussed in the next section.

**Mental Closeness**

As described by Mjöberg (2011) mental closeness is a mode of relating that can occur between a person and his contemporary, for example while writing or reading letters. Considering mental closeness is a promising way to discuss closeness in online chat togetherness, because—unlike joint involvement—it is not affected by the immediacy afforded by the environment but
rather works on an understanding the actor already has of the other. While text messages sent with digital communication technologies will typically reach the receiver much quicker than the physical letters in Mjöberg’s example do—text-based digitally mediated interaction, such as e-mail and instant messaging, is also asynchronous. Even in online chat interaction, like instant messaging, where the exchange of information can happen at a very fast rate, there is plenty of room for silence and reflection. Mental closeness could potentially be something that serves to fill those silences and help actors maintain their sense of being close to the other—despite a spatiotemporal distance from him or her.

One could say that the way in which the actor relates to her contemporary in mental closeness is through the proxy of her mental image of him or her. When the actor is writing a message to the contemporary, she is choosing her words, the tone, and the contents based on whatever knowledge is making up this image. Similarly, the person on the receiving end will read and interpret the message based on his mental image of the first person, and he will also get an idea of the way the first person was relating to him (or to the first person’s mental image of him) while writing the letter. As proxies, the mental images actors have of one another may allow them to feel that they share a connection, and perhaps even an involvement, in the exchange of messages, even when they are not jointly involved in the immediate moment. That mental closeness is indeed an element that helps spatiotemporal actors feel more closely together in digitally mediated environments is also supported by Jurkane-Hobein’s (2015) study of relationship maintenance in long-distance intimate relationships. Jurkane-Hobein found that a key factor for actors to maintain a sense of closeness to the other was their ability to vividly imagine him or her.

If the mental image of the other serves as a proxy for mental closeness, it follows that the actor must have an understanding of who the other is in order to relate to him in this way. While Mjöberg does not explicate this matter, the examples she uses imply that the actor has a characterological understanding of the contemporary. That would mean that the actor understands the other not as a social type (i.e. as a habitual contemporary) but as a specific individual. According to Schütz, characterological knowledge of another person is usually gathered through face-to-face interaction with this person. Thereby, what history of interaction the actors have had determines how they will characterologically understand each other when they are apart. Depending on this history, each of the characterological understandings an actor has of his current contemporaries would be very different—the checkout clerk at his local supermarket is a person he ‘knows’ differently than his mother, even though both might at present be characterological contemporaries to him. It is perhaps possible that the differences do not come down only to how much information he has previously gathered about these two contemporaries, but that there is also an emotional element involved—
compared to the check-out clerk, the actor’s mother is probably more important to him emotionally, and cannot be replaced for someone else.

Schütz also suggests that characterological knowledge is not always acquired through first-hand experience of the other in face-to-face interaction, but that it can also be gained from interacting with someone who knows the absent person characterologically. That Schütz points to hearsay as a sufficient source of knowledge suggests that he does not, in fact, believe that we can only apprehend distant others characterologically if we have previously spent time in their immediate physical presence, but that this happens as soon as we believe that this person is a unique individual that is closer to our reach than other contemporaries are. It is not, therefore, the immediate experience of the other that matters, but more likely the form our understanding of the contemporary takes (e.g. as an individual compared to a type) and the social distance between us (compare e.g. a neighbour to the Queen of England). These considerations would imply that it is possible to form a characterological understanding of someone one has only ever experienced in text-to-text interaction, which resonates with how actors do in fact meet and establish relationships with others solely through text-to-text interactions.

A question that emerges where mental closeness is concerned is what role emotions, intentions and also the understanding of the current relationship play. For example, a person who separates from her partner of many years, and moves in with someone she recently met, might in future text-to-text interactions feel more mentally close to her new partner than she does to her ex-partner. So even though she is likely to have more personal knowledge of her ex-partner, because she has known him for a longer time than she has her new partner, her attitude towards and understanding of each relationship could, in this case, potentially be more influential to mental closeness than the length of each relationship. As it concerns the way in which actors relate not just to the interaction (as in involvement) or to the other (as in mental closeness), but to the relationship in itself, this brings us to the third and final mode of relating to be discussed—namely Simmel’s idea of intimacy.

**Intimacy**

Simmel (1908/1950) describes intimacy and its counterpart strangeness (1908/1971) as forms of relating that solely occur within the confines of an encounter. However, while the one mode of relating (intimacy) is characterized as something that draws actors closer together, the other mode (strangeness) is seen as something that maintains a distance between co-present actors. Yet the actual behaviour in both modes can be identical, even in terms of how much personal and emotional information the actors exchange. So, what determines whether a situation in which actors are sharing private thoughts with each other is ‘strange’ or ‘intimate’?

Simmel’s discussion of strangeness suggests that actors frame the situation in a generalized way, which means that actors are not truly relating to
each other as unique individuals, and neither are they treating the relationship as one that has a future of significance. In that sense, strangeness comes across as a mode of relating that involves little risk of being rejected or hurt, as actors are not entirely invested in the relationship. Perhaps one can say that, in strangeness, actors are taking on an attitude of indifference. By contrast, intimacy is a form of relating that is entirely directed towards the unique and exclusive qualities of the relationship in question, and where the specific other and specific relationship matters to the actor. It stands to reason that whatever actors do or say in intimate situations like these—particularly when they disclose private information—always involves a considerable risk of disapproval and rejection; yet being willing to take those risks might also be what will help the relationship become close and intimate rather than distanced and ‘strange’. The difference between Simmel’s intimate and strange modes of relating could thereby be said to come down to the (indifferent or caring) attitude of actors, which in turn is shaped by how actors understand the frame. This would include how they understand their relationship, and the risks and benefits associated with certain activities.

The intimate, or caring, attitude emerges as a form of emotional investment that is uniquely tied to the actor’s relationship to a specific other. It follows that if actors are to share an understanding of the encounter as an intimate one, each participant’s intimate attitude must be successfully communicated or expressed. Given that the digitally mediated environment alters how information can be transmitted, it would appear that, for this reason only, these environments complicate intimate relating. However, the way Simmel describes how intimate and strange expressions and communications can be identical also suggests that this uncertainty could be present in all relationships under all interactional circumstances. So in both online chat encounters and in face-to-face encounters, there is a risk that the other is not relating to the relationship as intimately as the actor himself does. But how, then, do actors reassure themselves that the other is also framing the situation as intimate (or strange)?

According to Simmel, three elements appear to distinguish intimacy from strangeness: exclusivity, emotional investment (which would increase the perceived risks), and the relating as one between unique individuals rather than interchangeable types. In online chat situations, I would argue that the last element has already been captured by the concept of mental closeness. Mental closeness could be communicated through the choices actors make in conversation that in some way express that the actor has made these specific choices with the other person in mind. For example, the actor can communicate mental closeness by making a reference to something that the other has previously said or to a shared experience in the past, by initiating certain topics and not others, or by choosing words and phrases that bear special meaning to the relationship.
While exclusivity is technically decided outside of the ongoing encounter and relationship by the actions or attitudes that are not taken or held, certain behaviours in an encounter can also signal, or symbolize, exclusivity. To take a straightforward example, actors who experience a relationship as unique often verbally proclaim this to the other (e.g. ‘you are the only one who understands me;’ ‘you are my one true love;’ ‘we share a special connection’). Proclamations like these can of course be mediated in both face-to-face and text-to-text interaction, and may come across as either sincere and true, or empty and false, in both types of situation. That is because the validity of a statement such as ‘I love you more than anything’ is typically not decided in the current moment, but by all activities concerning the relationship that have preceded this statement of the moment. While the relative lack of expressed cues in text-to-text interactions may cause more doubt as to whether the professed exclusivity of a relationship is truly sincere, this type of uncertainty also famously exists in relationships that are predominantly maintained in face-to-face interaction—jealousy did not emerge as an effect of the Internet, after all.

As already mentioned, exclusivity is not primarily made up of a verbal agreement, but is rather composed of the actors’ experiences of taking a privileged and singular part in each other’s lives. Besides verbal statements, another way in which a sense of certainty about exclusivity comes about is through knowing each other’s everyday trajectories. This may be hard to maintain at a spatiotemporal distance, particularly for a partner that is entirely unfamiliar with the other’s spatiotemporal location as the ability to share, or even imaginatively share, his point of view is diminished.

For a couple that co-habits, yet spends most waking hours apart, it is for example easy for one partner to imagine what the other person is experiencing throughout the day, because he probably has his own first-hand experiences of the places she will be in and of the things she will see. He can imagine her approaching her office building, being caught in that day’s rainfall; he can also picture the restaurant she orders lunch from, and what she sees on her way home. However, he will struggle to imagine in the same detail how she spends her day and what she experiences when she is away on a business trip in a city he has not been to, even if she were constantly sending him updates about her day through instant messaging. The access the actor has to his partner’s daily trajectory is here fragmented, and thereby the exclusive and privileged standpoint from which he normally relates to her is somewhat threatened.

The same sort of imaginary vacuum would be common in digitally mediated relationships where actors are located in different parts of the world, or even just in different cities in the same country, where one or both of the participants cannot relate to the other person’s complete point of view. Similarly, it is clear that, in online chat environments, the actor is normally partly blind to what the other is experiencing in the moment, as the physical envi-
environments that others are located in will usually remain hidden from her direct experience.

While sharing someone’s point of view is not the same as experiencing things exactly from his point of view, actors will have a harder time imagining what the other person is experiencing the more spatiotemporally distant they are from each other. And so, in the online chat context, ‘exclusivity’ can rarely stand for sharing in each other’s entire range of possible experience, unless the actors who share an online chat situation are also co-located in a physical environment. Another rather important implication of this is that an actor who feels that she is currently in an exclusive (dyadic) online chat conversation might be unpleasantly surprised to learn that the other is simultaneously talking to (and perhaps even sharing what the actor is writing) with a physically co-located friend.

Interestingly, Beneito-Montagut (2015) also found that, although all of his informants would regularly maintain several text-based online encounters at once, they nevertheless felt disappointed if they realized that their interaction partner was doing the same. Of course, the ‘multi-interacting’ actor can still be perfectly clear about which encounter is the most intimate or exclusive even if the other person in that exclusive encounter may interpret the behaviour as a sign of the opposite. So while the overall exclusivity of a relationship may be equally hard to determine in any environment, there is a big difference between face-to-face and online chat environments regarding how certain actors can be about the exclusivity of an ongoing encounter.

The last element of Simmel’s intimacy concerns emotional investment, and of particular interest is how actors can communicate or express this investment, for example through signalling to the other that they are taking a personal risk. As has been explained, actors have more control over what they communicate in online chat environments than in face-to-face environments, which suggests that the perception of vulnerabilities and risks differs between the two. Of course, even if the environment and its affordances remain the same, different situations might carry different perceived risks depending on how the actors understand those situations. Yet, the increased control, particularly over expressed information, that is afforded by online chat environments suggests that there should be more general differences between how actors understand risks and their own and (other’s) vulnerabilities in face-to-face situations compared to online chat situations.

The behaviour of volunteering personal information, such as thoughts, experiences, or emotions that are not normally revealed to others, relates to both risk-taking and intimacy/strangeness in relationships. Therefore, the tendency of actors to be more willing to disclose personal information in online chat situations (and other asynchronous digital situations) than they are in face-to-face situations could be interpreted as a reflection of changing conditions for intimate (and strange) relating. This behaviour has been described by some as the ‘online disinhibition effect’ and explained as a phe-
nomenon that springs both from the actor’s anonymity in online environments, and from the way the actor relates not quite to the other but to his abstract mental image of him or her (Suler 2004). In other words, disclosure caused by the online disinhibition effect is characterized as a ‘strange’ mode of relating as it is thought to involve little personal risk-taking. Thus it would maintain actors at a distance, while they are at the same time being very personal with each other.

But it has also been observed that self-disclosure and emotionality are common in text-based situations, such as on online dating sites, where actors are seeking to establish personal or even intimate relationships (e.g. Rosen et al. 2008). This suggests that the act of disclosing private information cannot solely be explained as something actors only do in text-to-text interaction because they believe that nothing is at stake. In online dating, for example, the risk of being rejected should be a concern to most participants, which would suggest that, here, the sharing of intimate information is just that—a form of intimate, rather than strange, relating.

In a study on how actors express their inner feelings (or their ‘true selves’) in online discussion forums, McKenna and colleagues (2002) similarly found that such self-disclosing acts improved the chances that actors would establish close and durable social relationships. It has also been observed that intimate relating—in the sense of verbally disclosing feelings and private information—is not only more common in text-based digital situations, but that digitally mediated disclosures are also perceived as more intimate compared to face-to-face situations; moreover, the norms of reciprocity are also stronger here (Jiang et al. 2013). As Gibbs and colleagues (2011) point out, the norm of reciprocity also means that acts of self-disclosure can serve as an uncertainty reduction strategy: the more an actor discloses, the more she can expect the other to spill personal information. While this behaviour can appear overly strategic and non-intimate in nature, a reduction in uncertainty typically increases trust between actors, and trust is in turn linked to intimacy.

Self-disclosure could thus signal that actors are relating to the textually mediated relationship in an intimate way, but it can equally well indicate a relationship of strangeness. Whether a relationship is perceived as intimate rather than strange would come down to how the overall situation is framed (e.g. compare the frame applied to a situation on a dating platform to that applied to a platform for business transactions), the emotions invested, and also how successfully these emotional intentions are communicated and reciprocated between the participants in question. In lack of embodied expressions, a reciprocal sharing of private and otherwise normally concealed thoughts and feelings is perhaps the most straightforward way in which actors can communicate an intimate attitude, and maintain intimacy, in the online chat environment.
As described above, the lack of embodied expressions and the comparatively low immediacy in text-to-text interaction brings about new uncertainties, for example about what to expect of the other, what his spontaneous reactions are, or whether he is intensely invested and involved in the encounter or only marginally so. Relating to someone intimately in the online chat situation, and understanding this intimacy as something that is shared by the other, would therefore have to involve different risks and rely on different types of information compared to the face-to-face situation. I have here discussed how intimacy can be broken down into three elements.

First, intimacy involves a sense of shared mental closeness where the actor is directed to the other as a specific individual, and the other to the actor. How this is related to the affordances of the environment was discussed previously.

Second, intimacy involves a sense of exclusivity, which is linked to a sense of control of the other’s intentions, emotions, and current engagements. While the extent to which one’s own way of relating to the other is exclusive should be rather clear to the actor, it is always difficult to reach the same certainty regarding the other’s exclusivity. This uncertainty should be present in face-to-face situations and online chat situations alike. However, some conditions of text-to-text interaction—such as a spatiotemporal distance between actors—might influence how certain the actors are about how exclusive their relationship is.

Last, intimacy also concerns the extent to which one is emotionally invested in the relationship, and the extent to which one understands that investment as a reciprocated one. In regard to online chat environments, the question becomes how actors—despite the relative lack of (embodied) expressed information—manage to communicate that they are taking such a leap of faith, and how they reassure the other that the risk they have taken in so doing has been well received. Self-disclosing behaviour in online chat situations could potentially function to communicate investment in the relationship, and one way of invoking a feeling of trust in the other would be to reciprocate this person’s self-disclosures. Therefore, an actor’s sense of reduced risks in online chat situations might not be the only reason for the comparatively high disclosure of personal information, but instead self-disclosure could be an efficient way of being closer together in situations where the physical bodies of actors are absent. This would mean that the behaviour of volunteering personal information could be a way of compensating for spatiotemporal distance, so that intimacy can be achieved.

Chapter Summary

In this chapter, I have discussed togetherness in digitally mediated environments—specifically online chat environments—from the perspective of co-
location and from the perspective of co-presence. I began by discussing how, in social presence theory, togetherness is understood from the perspective of co-location, meaning it is focused on mere co-existence, and the recognition of such co-existence, in environments. There are several limitations to this approach, and I have argued that these limitations over-simplify the matter of togetherness, at least in regard to togetherness in online chat environments.

For example, the social presence theories emphasize the human tendency to react socially to stimuli, which is something of a non-issue in digitally mediated environments generally (because these environments also afford social stimuli). Moreover, the approach portrays togetherness as little more than an instinctual reaction, while failing to capture how actors themselves actively and consciously contribute to the making and maintenance of togetherness. Another line of thought suggests that actors experience social presence (or togetherness) when they experience the digitally mediated environment as if it was a physical environment. The related idea of togetherness as an ‘illusion of non-mediation’ is problematic because it presupposes that the environment always goes unnoticed in face-to-face interaction, or that actors are not capable of seeing their environments for what they are without at the same time losing their ability to engage with others. I argued that for online chat environments, actors are unlikely to experience that the text-to-text interaction is identical to face-to-face interaction, or that they are not interacting by means of a technological medium. Nonetheless, the popularity of text-to-text interaction suggests that these interactions are still socially rewarding to actors, which suggests that these technologies do afford a sense of togetherness.

That co-location only captures a small part of what is entailed by togetherness was further illustrated by the example of people’s text-to-text interactions with VCAs. Here, I concluded that the reason VCAs so often fail to successfully engage with humans in text-to-text interaction is not that they cannot be co-located with humans in these environments, but that they struggle to be successfully co-present with them. In a sense, the limitations of the co-located perspective on togetherness come down to the fact that it can only explain how actors come to understand that they are there together, not their understandings of with whom they are together or how they are being together.

In the second part of the chapter I focused on togetherness as co-presence. Drawing on a number of microsociological observations of how actors are together and get closer together in face-to-face situations, I discussed how these observations and concepts would translate to online chat situations.

First I examined how the affordances of online chat environments affect how actors come to understand the situation. The way in which online chat environments allow actors to communicate and express information is close-
ly linked to how they reach an understanding of the situation, and in effect to the expectations they form about their own and others’ behaviour in it.

In particular, the expectations actors have in the situation are connected to the level of control they have over the information they transmit about themselves. In contrast to face-to-face environments, expressed embodied information is usually reduced to a minimum in online chat environments, which limits the access to information that participants involuntarily give off. Therefore, information of certain types—such as information about category belongings and about spontaneous and immediate reactions—loses some its value as more true and trustworthy information about the other in the online chat situation (because this information is no longer given off, but consciously given). In addition, environments that do not afford spatiotemporally immediate interaction also make it more difficult for actors to take the other’s point of view. However, since the loss of expressed and contextual information goes both ways, actors may sense that they have more control over what impressions and expectations others form of them, and at the same time feel less certain about what to expect of others.

Drawing on the concepts of involvement, mental closeness and intimacy, I continued to discuss distance and closeness between co-present actors in online chat environments. As the focus was closeness in togetherness, and not the isolated participants’ feelings of closeness in the situation, I was primarily concerned with how involvement, mental closeness, and intimacy are communicated and shared between actors. Involvement proved to be a difficult concept to match with the lack of spatiotemporal immediacy of online chat situations. While actors can of course be as involved in text-to-text interactions as they are in face-to-face interaction, the problem that emerges is how such an involvement is made visible to other participants.

If involvement is to be a mode of relating that works to draw actors closer together in the situation, then actors must perceive their involvement as reciprocated by the other. But particularly when it comes to the more intense forms of involvement described in the literature, such as joint spontaneous involvement, a non-reflected flow of spontaneous contributions does not appear to be compatible with true involvement in an ongoing text-to-text encounter. However, involvement could potentially be communicated indirectly, for example through messages that convey that the sender has put plenty of thought and effort into what is written (a similar line of thought is captured in Walther’s 2007 theory of hyperpersonal interaction). But involvement would then be the act of making qualitative contributions to the interaction, which is a conception that has strayed rather far away from what involvement originally stands for in the microsociological literature.

The concept of mental closeness captures the idea of making qualitative contributions to written conversation, but it also describes how spatiotemporally distant actors can still be directed to each other as specific individuals, and thereby experience a certain closeness. The way in which qualitative
contributions enter into the picture is through the act of composing a letter or a written message while being directed to the mental image of the receiver. Each word and sentence is thus chosen with this particular other in mind, and there is little in the way of communicating with the other in a generalized sense. Here, the knowledge an actor already has of the distant other should make a difference for the extent to which this other feels that the received message was intended for him and no one else. The more an actor experiences that he, specifically, is targeted by a message, and the more he experiences that he is targeting a unique other, the ‘closer’ to the other he should feel. I believe that mental closeness is likely to be an important aspect of how closely together actors feel they are in text-to-text interaction, and that it would be a relevant aspect to further explore in empirical studies of closeness and distance in online chat encounters.

I finally discussed intimacy as a mode of relating that could be said to bring actors closer together. In online chat situations, I argued, intimacy must always involve some form of mental closeness, because otherwise the relating would be ‘strange’. But involvement also concerns (perceived, if not actual) exclusivity and a sense of emotional investment in one’s relationship to a specific other. Due to the reduced control over the other person’s involvements and spontaneous reactions in online chat interaction, it is possible that an actor will experience less certainty as to whether his perception of the relationship as an exclusive one is truly shared with the other. Yet, no matter how a relationship is mediated, this uncertainty is present in most (if not all) relationships that are believed to be exclusive in some way.

Emotional investment was the last element of intimacy to be discussed, and I suggested that communicating such an investment largely comes down to how one communicates to the other that a personal risk is being taken. As has earlier been discussed, however, the way in which actors can control the information they send out to others in online chat situations implies that the perception of what behaviour involves a personal risk would differ from how actors understand risks in face-to-face situations. I discussed how self-disclosure could potentially be a way for an actor to communicate that a leap of faith has been taken, and that she trusts the other enough to willingly make herself vulnerable in the situation.

While the theoretical nature of this study of course prohibits me from drawing any conclusions about how actors de facto come to be together in the online chat environment, or if and how they can experience greater or lesser closeness to the other in such togetherness, I find the perspective of co-present togetherness to be more promising for future empirical studies than I do the approach suggested in social presence theory. I hope to have made a compelling argument as to why togetherness must happen in framed situations rather than in the shared informational situation, and to have shown how a microsociological perspective on togetherness might provide a way forward. While some of the microsociological concepts discussed here,
such as joint spontaneous involvement, might be of limited relevance to the understanding of how actors come closer or more distantly together in the online chat situation, I believe that the concepts of mental closeness and (Simmel’s definition of) intimacy are worth investigating further. Other aspects of situated behaviour, such as interaction rituals and social emotions, could potentially have been brought forward here, and the topics of self-presentation, embodiment, and the risk/trust dynamic in text-to-text interaction could have been discussed in greater detail. However, despite its limitations, I believe that the discussion of co-presence has shown both how the value of microsociological theorizing on togetherness extends beyond the face-to-face situation, and also that there is a need to further develop and appropriate these microsociological models and concepts if they are to be made fully suitable for the new conditions of social interaction.
5. Summary and Conclusions

This thesis set out to study togetherness in digitally mediated interaction. More specifically, the purpose was to examine theoretical understandings of togetherness both within the field of social presence theory and the field of microsociology. The motivational ground for this endeavour is that, both in the media-theoretical and the microsociological literature, the theoretical conceptualizations of what digitally mediated togetherness represents are insufficient, at the same time as a significant number of our everyday social interactions have migrated to digital environments. Micro-level analysis of mediated interactions is thus highly warranted, and along with this comes a need to develop theoretical and analytical resources that can aid such research. This thesis’ focus on togetherness is of course merely one out of many areas in the study of mediated social life that may require conceptual and theoretical improvement (as the field of study is still young); however, I believe that the concept of togetherness is of particular relevance because of the elementary character of the phenomenon it denotes and its central place in the broader understanding of everyday social life.

Digitally mediated social arenas come in many different shapes and forms. To name a few, users can engage with others on social networking sites, online forums, weblogs, virtual worlds, instant messaging platforms, massively multiplayer online games, and online chat rooms. Taken together, the means of communication afforded in these arenas include text-to-text transmission, avatar-to-avatar interaction, video chat, and phone calls. In addition, the hardware required to take part in digitally mediated social life varies depending on the arena of choice: some are more suitable for smartphone access (e.g. the instant messaging service WhatsApp), others are better accustomed to personal computers (e.g. MMORPGs like World of Warcraft), and yet others can easily be used on any type of hardware (e.g. Google’s instant messaging service Hangouts or the social network platform Facebook).

Given the heterogeneity of digital social life, this thesis has narrowed the focus to primarily discuss certain forms of text-to-text interaction. More specifically, it has focused on text-to-text interaction that users tend to treat as synchronous interaction (despite the fact that text-to-text technologies can only afford near-synchronicity), which comes down to different types of online chat applications (e.g. instant messaging services and online chat rooms). Even though this is not primarily a thesis about online chat interac-
tion but a thesis about digitally mediated togetherness, I have, throughout the thesis, attempted to tie the theoretical discussions about togetherness to the more concrete example of online chat interaction.

In the introductory chapter I presented a number of questions that awaited answers. I will now review each of these questions and the answers I have reached. The first question asked why the conceptualization of togetherness has become problematic since the advent of electronic and digital communication technologies. In line with Meyrowitz’ (1985) discussion of how electronic media separate social place from physical space, I approached this question as one that does not firstly concern togetherness *as such*, but as one that concerns the conceptualization of *where* actors must, can, or will be together. In Chapter 2, therefore, I examined understandings of what constitutes the where of togetherness and the implications that new forms of interactive technologies have on these understandings. In addition, Chapter 2 also provides a number of possible answers to my second question, namely the question of where actors are together, specifically when they are together in online chat.

In sociology, the *shared physical environment* has traditionally been considered the where in which actors can meet and interact. More precisely, actors are described to meet and interact in *social situations* that are in turn considered to be bound to shared physical environments. That is because these situations are described as instances where actors can perceive each other immediately and with their naked senses, which is a type of spatiotemporal immediacy that is exclusive to the shared physical environment. However, the behaviour in these physically anchored social situations is not merely structured by the environment (e.g. by how the environment affords actors the ability to perceive others and be perceived by others) but also by socially constructed norms and regulations.

Behaviour in any single physical environment within which actors are socially situated can be governed by very different social frames (as Goffman would put it). In turn, a frame can be roughly divided into more generalized structures of behaviour, such as the larger cultural frame or a frame that is general to a type of place (e.g. airports, public transports, or schools) and structures that are more specific to the circumstances of a particular meeting. For instance, actors can enter into a social situation at a certain physical location (such as a church) in order to attend a specific occasion (such as a wedding) and during this occasion they may hold several different conversations with different participants. In this example, an actor’s behaviour is structured by rather general frames (e.g. culture, and by the frames that are ordinarily applied to churches and weddings), but every time she goes from interacting with one participant (say another wedding guest) to interact with another (e.g. a waitress or the bride) she will also change her behaviour, meaning that more specific (here: role specific) frames are also in use.
The quest of establishing where actors are being together is now beginning to look like the neat reassembling of a Matryoshka doll: two actors can be together in a specific frame within a general frame within a social situation within a physical environment. However, when electronic and digital communication technologies enter into the picture, this neatness starts to fall apart. That is because the way in which these technologies allow spatially distant actors to interact in near immediacy challenges the assumption that actors can only share a situation when they are also in each other’s immediate physical presence. Suddenly, it would seem, the Matryoshka doll can be complete without the layer representing a shared physical environment—or can it? Are two people interacting through an online chat service together, or are they not?

To further complicate the matter, the new social arenas afforded by digital communication technologies are frequently (if not mostly) portrayed with the same language used to describe physical environments. For example, websites that offer online chat are described as places containing one or several ‘rooms’ (i.e. chat rooms) in which users can ‘hang out’ and engage in conversations with other users. In other words, digitally mediated arenas—particularly those arenas that provide some durable structure, as often found on the World Wide Web—are (by scholars, designers, and users alike) conceptualized as spatial environments that actors can ‘visit’ or ‘inhabit’ with their digital ‘bodies’. There are some very good reasons for this, and a number of problematic side-effects.

To begin on a positive note, the use of spatial metaphors helps us visualize and make sense of what is essentially a very abstract phenomenon. Designers of user interfaces find it useful to employ these metaphors to make it easier for users to intuitively understand how to operate the system, as users can extrapolate from their already existing understanding of the spatial environment. But on the negative side, this practice confuses the understanding of where we are, because it invites us to think that we are doing the impossible, namely being located in two spatial environments at once. Yet at the very same time, we tend to be aware that our bodies remain firmly rooted in only one spatial environment—the ‘real’ one.

This last sentence points to another problem with the spatial description of our digitally mediated whereabouts. Since we are still aware that the ‘virtually spatial environment’ is not actually a true spatial environment, we often perceive these environments as inferior, and less genuine, imitations of the real thing. In effect, anything that happens in the digitally mediated arena—such as social interactions and the emergence of social relationships—risks being perceived in the same way (i.e. as unreal and artificial social substitutes).

If the where of being there together follows this example, it would seem that being in, for example, a chat room together should signify the virtual colocation of virtual bodies in a virtually spatial chat room, in what mimics,
but is nevertheless entirely opposed to, how our physical bodies are physically co-located in physical environments. But digitally mediated arenas are not (and can never be) exactly the same as physical environments, even though they are, at the end of the day, intrinsically part of and co-evolving with the material world (Graham 1998, see also Latour 2005). Following this, there is no such thing as environments separated from the physical world, and so being in a chat room would in no way indicate a simultaneous separation from the physical environment. And just because we are in some way making use of media in our physical environment, that does not mean we are losing our grip on reality, because everything that exists is part of our reality. Therefore, one last way of looking at mediated arenas is not in terms of free-floating environments that we can only access by in some way exiting all the other environments we are in (which is what we must do if we want to move from one physical environment to another). Instead, a mediated arena would be embedded in the environment that our physical bodies are located in, and in that way it alters the affordances of the overall environment. More specifically, it increases the perceptual and interactive affordances offered.

However, Meyrowitz (1999) still prefers to describe mediated arenas in terms of environments, which is also how I have referred to digitally mediated arenas (such as online chat rooms and instant messaging platforms) in this thesis. That is because, when the spatial and material aspects specifically associated with physical environments are bracketed, physical and mediated environments are alike in that they afford overarching possibilities and constraints to what and how information can flow between actors. While Meyrowitz’ (1985) previous conceptualization of the social situation as a specific pattern of information flow has not been considered in this thesis, on the grounds that it is so wide that it loses some of its analytic purpose (for example it implies that actors can be equally co-present with movie stars on the television as they can be with the person with whom they are having a face-to-face conversation), I do endorse his depiction of mediated arenas as (informational) environments.

In brief, Chapter 2 concluded that the underlying reason why togetherness has become more difficult to conceptualize is that electronic and digital communication technologies separate our social whereabouts from our physical ones. In addition, the portrayal of digitally mediated environments as spatial environments is unfortunate, because it gives the impression that concepts developed for face-to-face togetherness can be directly applied to digitally mediated togetherness. But the shoe does not fit, and at worst such attempts may entrench the idea that digitally mediated togetherness is not entirely ‘real’.

As for the second question, the chapter surveyed many possible wherees in which actors can be together, and these are captured in the following three categories:
(1) The (informational) environment. Being jointly situated, or co-located, in an informational environment involves sharing the same unique set of environmental affordances for how and what type of information can flow. This does not require any established relation of reciprocity between actors; however, the sharing of an informational environment is a requisite for the two categories that follows. In addition, from the perspective of the actor, the shared part of her informational situation does not need to seamlessly overlap with the totality of the particular informational situation she is currently in.

The informational environment of the instant messaging service WhatsApp would thus include the means of information transmission afforded (here: text messages and the sharing of photos and small audio and video files), but also information communicated to users through e.g. the interface design. Being ‘there’ together in the informational environment of WhatsApp would therefore denote a state in which two or more people have access to the affordances and the information provided by the application, yet where they do not necessarily have shared access to each other’s complete environments.

(2) The social situation. Being co-present in an unfocused social situation means being in reach of the other, where the other is also in reach of oneself. It is a situation of mutual monitoring possibilities, where actors are reciprocally aware of each other. Co-presence is here governed by a diffuse (unfocused) behavioural structure or frame.

When logged into the online chat room Zobe, the actor is constantly reminded of who else is logged into the environment through a list of active users that sits on the right hand side of the website interface. These users are within her social reach in the environment, as she is free to send a private message to whomever she wishes to converse with. In addition, users can also send messages in the public chat room, where the information will be made available to everyone (and sometimes no one in particular). While there is comparatively little regulation as to what a user can post in the public chat window, certain types of behaviours—such as spamming, flaming, or excessive use of capital letters—are generally considered faux pas in the online chat context (and are sometimes sanctioned). To share the unfocused situation of the online chat room thus requires users to be logged in and mutually accessible to each other, and that their behaviour is governed by some general rules of conduct.

(3) The focused social situation. In a focused social situation actors are reciprocally oriented towards each other, and possibly even overtly interacting. The conduct of actors is guided both by general and specific frames.

Two Zobe users that are engaged in a private chat conversation can thus be said to be together in a focused social situation. Here, the general rules mentioned above may still apply and, if not followed, sanctioned (for example, the user may terminate the conversation if the other starts spamming or
insists on ‘screaming’ her messages by using all capital letters), but the behaviour of users will also be guided by frames that are more specific to the situation.

The three categories are broad enough to accommodate both face-to-face meetings and digitally mediated ones. However, what may fall outside the scope of this categorization is the notion of being together in a place. The concept of place does carry connotations that may better capture the sentimental aspects (such as feelings of belonging in a place, or place attachment) of the where in which an actor is than the concept of the social situation does. The actor’s relationship to the place she is being in is a relationship that may influence her relation to the actors she is being with in that place. As Milligan (1998) suggests, there is a mutually influential relationship between place attachment and social interactions: social interactions within a place shape place attachment and in turn the place attachment creates expectations on—and therefore shapes—future interactions within the place. Place attachment can thus be seen as a social construction that in time begins to shape the frames that govern meetings within environments. While Milligan’s study is concerned with physical places, I believe it would be fruitful, in future research, to also explore the relationship between digital places, place attachment, and digitally mediated togetherness.

The third and fourth questions asked what togetherness means in two fields of study, and were considered in chapter 3. In the field of social presence theory, the conceptualizations of social presence (defined as ‘being there together’) fall within three broad categories, social presence can either be seen as 1) the existence of certain types or a certain number of social cues in the environment, 2) related to self-presentation, and 3) an experience of external events (and here usually the experience of others or of non-mediation). In general, one can say that most of these perspectives in one way or the other concern how the informational environment affords actors the ability to perceive others, or make themselves perceived by others, in the environment. With some exceptions, these conceptualizations also do not take into consideration how actors, and the social norms imposed on actors, contribute to social presence. Therefore, the extent to which there can be ‘more’ or ‘less’ togetherness between actors in this discourse is explained by variations in what the environment affords, where the ‘richness’ and ‘immediacy’ of the afforded information is particularly emphasized. One implication of this is that online chat environments will always be placed on the lower end of the togetherness scale, as the comparatively low immediacy and richness offered by this technology would, according to social presence theory, only afford online chat users a very limited experience of togetherness.

The second part of the chapter examined what being together could stand for in the microsociological literature, specifically in interactionist and phenomenological writings. Here, a picture of being together emerged that ex-
tended beyond actors’ (reciprocal or non-reciprocal) observations of each other, to also include the frames actors are obliged to act in accordance with in order to maintain togetherness. Depending on the frame in use, togetherness can be unfocused or focused. However, without a frame, actors cannot be together in this perspective—in fact a frame and togetherness go together inseparably. In addition, the frame is also understood as dependent on the affordance of spatiotemporal immediacy.

When it comes to the degree of togetherness, the microsociological literature agrees that how richly and immediately actors can experience each other affects togetherness. But while Schütz acknowledges that even in direct (spatiotemporally immediate) situations the concreteness with which actors experience each other can vary (depending e.g. on the point of view actors have of each other), the field makes a rather sharp distinction between ‘direct’ realms in which actors can be together, and ‘indirect’ realms in which they cannot be together. In spite of that I went on to examine how, in the indirect realm, the actor is described as experiencing certain contemporaries (i.e. characterological ones) as closer to her than others are, even if she is technically not currently there with any of her contemporaries. Drawing on Mjöberg (2011) I also described a particular form of relating to characterological contemporaries, namely mental closeness. I suggested that the concepts of characterological contemporaries and mental closeness would be useful in the analysis of togetherness in online chat, as online chat technologies can only afford near immediacy.

In the last sections of the chapter I discussed degrees of togetherness within spatiotemporally immediate social situations. Here I reviewed Goffman’s concept of involvement (and misinvolvement) and Simmel’s (1908) concept of intimacy (and strangeness), as they both refer to interactional (in the case of involvement) and relational (in the case of intimacy) elements that affect the closeness or distance between actors that are already together in a framed situation. Later on I also emphasized how actors cannot just practice involvement and intimacy in an encounter however they like, because the frames governing the situation will ordinarily regulate these elements.

Chapter 3 concluded that, in social presence theory, togetherness represents a number of different phenomena that all have to do with what social information the environment affords, and how it affords the transmission of this information. This is admittedly a very vague description of togetherness in social presence theory, but it does manage to encapsulate all the definitional categories identified. In this discourse, distance and closeness between actors who are being together is generally decided by informational variables: the richer and more immediate the information afforded, the closer together actors are.

In microsociology, being together represents a framed social situation in which actors are reciprocally aware of, and oftentimes also overtly respon-
sive to, each other. Distance and closeness in togetherness are here described as decided not primarily by the spatiotemporal conditions of the situation, but by interactional and relational variables that are in turn regulated by the social frame. Because of this I proposed that, even though they are originally intended for the analysis of face-to-face situations, the concepts might still be of value in the analysis of togetherness in online chat and other digitally mediated situations.

The selection of literature could of course have been extended to include other, both media-theoretical and microsociological, discourses that relate to the notion of togetherness. However, social presence theory is, to my knowledge, the only larger theoretical discourse that has made a substantial and focused effort to arrive at an understanding of digitally mediated togetherness. As for the microsociological theories discussed (and the many theories that were not discussed), the literature included in the review was chosen because it highlights and details some of the more elementary aspects or necessities of social meetings and the being together of actors. The survey of relational and interactional factors that increase or decrease actors’ togetherness in focused encounters could however have been extended, for example to include insights from the sociology of emotions. That task is, however, one that I must pass along to future studies.

The two remaining questions were dealt with in Chapter 4. First, I argued that the reviewed discourses are focusing on two different *wheres* in their conceptualization of actors being (somewhere) together. I did not mean this in the rather obvious sense that the one field is concerned with digitally mediated environments and the other with shared physical environments (which means that they are dealing with togetherness in different types of informational environments). Drawing on the three types of *wheres* outlined in Chapter 2, I instead argued that social presence theory is essentially a theory of actors’ co-location in informational environments, while the microsociological theories portray togetherness as something that happens between actors that are co-present in framed social situations. Recalling the characteristics of different types of *wheres* (listed above), the different approaches of social presence theory and microsociology should have noticeable consequences for how togetherness in online chat is understood. But given the emphasis on face-to-face interaction in the microsociological literature, what is (largely) absent from the study of digitally mediated togetherness is an approach that directly locates the phenomenon in framed social situations.

I then proceeded to discuss the two approaches (co-location and co-presence) in relation to meetings in online chat environments. First, I discussed togetherness as co-location and argued that this is a limited approach in that it only focuses on the relationship between the actor and the environment (and the information presented to her in it). I pointed out three lines of thought in this understanding of togetherness that I believe are particularly problematic when applied to online chat specifically, and digitally mediated
interaction more generally. These are firstly the portrayal of togetherness as an impulsive reaction that, secondly (and in effect), depicts the actor as a passive and debilitated receptor of social information. The third line of thought portrays being there together as the illusion of non-mediation (an idea that is proposed in many, but not all, conceptualizations of the phenomenon).

I argued that the first idea is not relevant to use in the analysis of togetherness in online chat or any other digitally mediated environments, as all such environments are equipped to transmit stimuli that can elicit a social reaction in humans. I also proposed that the second approach is limited because it suggests that the nature and quality of the social stimuli that reach an actor are what first and foremost determine how strongly she feels that she is being with others in an environment. The possibility of togetherness being something that actors are actively participating in and creating is thereby replaced by the idea that togetherness is determined by the informational richness of the environment.

This would imply that within any given environment, such as an online chat room, the ‘amount’ or ‘quality’ of the togetherness between actors is always constant, because the informational affordances of this environment remain constant. So for example, an actor who is logged in to an online chat room but not communicating anything in the environment is just as much ‘there’ with the other chatters as a user who is actively participating in the discussion unfolding in the public chat window. This is practically the same as saying that a person who sits hidden away in the corner of a cocktail party is just as much there with the other guests as a person who is engaged in the ongoing conversation. While these examples correctly suggest that both the ‘passive’ and ‘active’ actors are co-located with others in environments, I argued that, if mere co-location can in any way be called togetherness, it is certainly a very weak form of togetherness.

As for the third approach, I argued that it is unlikely that actors can only experience that they are there with others if they are led to believe that the mediated interaction is not mediated. Unlike more immersive mediated environments (e.g. virtual environments and virtual worlds), online chat environments are not designed to give the user the illusion that she is meeting others face-to-face. I proposed that the fact that actors keep using online chat technologies, keep returning to their favourite chat rooms, and keep spending a considerable amount of time in these types of text-to-text interactions at least strongly suggests that actors do experience that the way they are being together in these environments is real and significant. To further emphasize the limitations of the perspective of co-located togetherness, I showed that the approach cannot explain why VCAs so often fail to successfully be together with actors in online chat environments, and I suggested that these failures may be better explained from the perspective of co-presence.
Second, I discussed the perspective of co-present togetherness. As I had already argued in favour of a more situational perspective on togetherness, what remained was to see whether the concepts and models used in the study of face-to-face encounters would also be suitable in the context of digitally mediated social contact. In line with Rettie (2009), I firstly established that one may indeed think of certain forms of text-to-text contact (i.e. the forms that are treated as synchronous, such as online chat interactions) in terms of unfocused and focused encounters. Thereafter, I proceeded to discuss the implications of online chat environments on such framed situations.

I began with a discussion of frames and framework, and ended by discussing the three modes of relating that were, in Chapter 3, described as affecting the closeness and distance in encounters (i.e. involvement, mental closeness and intimacy). Here, I concluded that, due to its dependency on the degree of (actual) spatiotemporal immediacy, the value of using the concept of involvement (in the Goffmanian sense) in the analysis of the degree of closeness in online chat togetherness is limited. By contrast, I found that the concepts of mental closeness and intimacy had greater potential for the study of online chat togetherness, as they are less sensitive to the spatiotemporal immediacy afforded by environments.

So to the remaining question, the one that asked which perspective of togetherness, at least for the study of online chat togetherness, carries greater theoretical and analytical potential, my answer is the co-present perspective. That is because togetherness must be seen as something that occurs not in the intersection of environments and (passive) individuals, but in the much busier intersection of environments, social situations, frames, and actors. The study of co-present togetherness is the study of how people, when co-located, are ‘doing’ togetherness under a range of different structuring circumstances. It takes into account the environment’s influence on social situations and (in effect) on togetherness, but it is also concerned with how actors negotiate togetherness within situations structured by environmental, social, interactional, and individual affordances. Co-presence is thereby an arguably broader and more complex social phenomenon than co-location is, but it is potentially also truer to the questions of what it means to be somewhere together, and of how this phenomenon can best be theoretically conceptualized.

Since the discussion in Chapter 4 primarily revolves around one form of digitally mediated interaction (i.e. online chat interaction), the results are of course limited in scope. Nonetheless, I believe that the discussion takes some initial steps in the direction of a more general theory of digitally mediated togetherness, one that can account for the busy intersections of co-present togetherness. Of those steps, I consider the discussion and evaluation of how different modes of relating may serve to bring actors closer together in online chat encounters to be the most novel, and perhaps most important, contribution. The reason for this is that my discussion takes the understand-
ing of togetherness beyond the necessary conditions for the phenomenon to occur, while also showing that, even in the online chat situation, it is unlikely that distance and closeness are decided solely by environmental, specifically spatiotemporal, variables. However, for future studies, my list of interactional and relational modes of relating would benefit from the inclusion and evaluation of other social forms of relevance to the closeness and distance in digitally mediated togetherness. Also beneficial would be more empirically focused studies, as well as studies broadening the scope to include digital communication alternatives other than online chat services.

***

This dissertation sits on the border of two distinct fields of study: the field of media and communication studies, and the field of microsociology. I believe that the present work contributes to both of these fields. To the mediatheoretical subfield of social presence theory, my review and critique of the conceptualization of social presence identifies weaknesses in the theoretical constructs that should be tended to. In addition, I also offer an alternative conceptualization of togetherness, namely co-present togetherness, that (with the exception of the view of social presence as the illusion of non-mediation) does not contradict the field’s existing conceptualizations of togetherness as co-location, but rather extends it in a way that solves several of the identified problems in the theories.

As for the field of microsociology, this thesis contributes to the theoretical understanding of a part of everyday social life that has thus far been little explored in the field (particularly from a theoretical point of view), namely our digitally mediated social encounters. It does so by looking more closely at a subset of digital communication technologies, namely online chat technologies, and by discussing what togetherness would represent in online chat interaction. Furthermore, the thesis highlights that there is a need for microsociology to theoretically catch up with the recent, yet substantial, changes in how actors meet, interact, and establish relationships, since there have been few attempts to formulate a formal microsociological theory of digitally mediated interaction. While the work presented here is by no means a formal theory of digitally mediated interaction, the insights may be of use in what will hopefully be many future microsociological endeavours to establish such a theory.

Through the presented examination and analysis of the academic discussion of togetherness, I believe this dissertation has achieved its overarching, as well as its more specific goals. It has contributed to the growing academic understanding of what it means to be together in contemporary social life,
and it has also, if ever so slightly, narrowed the gap between media theory and microsociology.
References


42. Ulla Jergeby: Offentlig miljö som arena och kuliss. Att se, bli sedd och mötas på stadens offentliga och bostadsnära platser. 1996.
47. Peter Sohlberg: Mål och mening i samhället. Funktionalistiska program i samhällsvetenskapen. 1999.
63. Fereshteh Ahmadi (ed.): Coping with Cancer in Sweden – A Search for Meaning. 2015.