Master thesis on Cognitive Systems and Interactive Media

Universitat Pompeu Fabra

I know what you type: empathy in real-time typed Computer Mediated Communication

Micozzi Ferri, Lorenzo

Supervisor: López-Carral, Héctor

Host Research Group: SPECS

July 2022

Contents

1 Introduction **1**

1.1 Problem statement 1

1.2 State of the art 2

1.2.1 Affective Computing and its insights for CMC 2

1.2.2 Social and cognitive interaction in the digital era 3

1.2.3 Seeing others is understanding others: Mirror Neurons and Empathy . 4

1.2.4 Possible interventions: Experimental and Critical Design for Well-being 5

2 Methods **7**

2.1 Hypothesis and General Methodology 7

2.2 Experimental design and set-up 8

2.3 Procedures used to obtain data and results 9

2.3.1 Subjects 9

2.3.2 Data collection and analysis 9

3 Results **12**

3.1 Ability to identify internal state of the other 12

3.2 Perceived interpersonal closeness 14

3.3 Revisions and coordination during the dialogue 14

3.4 Synchronization of typing patterns 15

4 Discussion **16**

4.1 Discussion 16

4.2 Conclusions 17

List of Figures **19**

Bibliography **20**

Abstract

We tend to give the design of the environments that host our daily digital interactions with other humans for granted. As we march towards a future where our digital life is always more prominent, is it time that we start asking more questions about the type of mediation that the digital tools we use operate on our social, emotional and cognitive experiences?

In this thesis I will try to highlight the importance of researching experimental digital environment and tools that take into account the nature of human communication and interaction.

As a stepping stone towards a clearer understanding of what we have access to, in terms of empathic relation with the "digital other", right now (and what could we aspire to in the future), this research leverages the emerging knowledge about keystrokes dynamics and typing patterns as a window on the dyadic dynamics that are observed during the experiment conducted.

Given the complex nature of language and communication, observations based on affective computing research are framed in a social, neuroscientific and cultural context, as to give a complete map of the issue.

While this research focuses on experimental text-based communication, I hope that many will be inspired to conduct their own studies, proposing new ways of being with the other in the digital era.

Keywords: Keystrokes dynamics; Typing patterns; Digital well-being; Social neuroscience;

Chapter 1

Introduction

1.1 Problem statement

Research in affective computing keeps widening our understanding of how significant our interactions with and through digital means are. In the perspective of human-computer interaction, systems that estimate our internal states by analyzing our interaction cues could increasingly help us build applications that adapt to human emotional feedback. Similarly, digital phenotypes can prove to be an important tool in forecasting and preventing mental health issues.

If these interaction cues can reveal so much about ourselves, could they also have an important role in a human-human digital interaction? Can a person infer internal states of others by observing their interaction cues? (The same cues, for example, that we expose to affective computing systems, to allow them to behave accordingly). Would exposing more of these cues in human to human interactions improve the effect that digital environments have on our wellbeing? Would it make us more empathic? Would it improve our performance on specific tasks?

Finding answers to these questions seems to me vital for designing technology, applications, platforms and environments through which we healthily foster our relationships as affective human beings.

Much of what we are online can sometimes be a simplistic and, in the scope of our interactions with others, alienating representation of us. Could exposing more of our "implicit interaction cues" to others counteract this tendency? For example, could humans be able to infer internal states of others by observing their typing patterns? If so, following the typing example, would the ability to see someone else compose in real-time their "text" (eg. in a chat / collaborative text / etc.) influence (possibly positively enhance?) their understanding of the others?

Another very interesting perspective on this matter regards the long term effects of different digital applications on our emotional and cognitive abilities. Do we absorb and adhere to the standards of digital interactions? How do they influence the development of our emotional and cognitive maps? Does using a "non-affective" computing system make us develop "non-affective" behaviours (inside and outside the digital environment)?

1.2 State of the art

1.2.1 Affective Computing and its insights for CMC

Research in the area of affective computing has attracted a lot of attention to the possibility of formulating models able to predict internal states of someone, based on the analysis of their patterns of interaction. In an era where a massive part of our communications with other humans has moved to digital platforms, more and more researches are interested in understanding what kind of information can be extrapolated by observing how we use the technology that gives access to these channels.

While a very broad variety of digital input dynamics has been observed to bear great significance on what can be understood about the person "behind the screen"[1], the present research is going to focus on real-time typed computer-mediated communication. Latest research suggest that typing patterns and keystrokes dynamics are very rich form of input to analyze: in recent studies, keystrokes features like "flight time" (the time passed between two key presses) correlated with valence and arousal

of images (selected from Open Affective Standardized Image Set) shown to participants asked to describe them [2]. Typing patterns were also observed to change based on the utterance's functional purpose and self-assessed rapport between two speakers communicating in a chat [3].

Being able to observe correlations between keystrokes dynamics and internal states and even predict internal states based on the features of the typing patterns, lay the scientific basis for considering these features an important gateway for the understanding and of the others in a digital environment. But how an why should we leverage this information?

1.2.2 Social and cognitive interaction in the digital era

"The advent of interconnected smartphones, the parallel development of social media and of the related subsidiary applications, have literally shifted the balance towards mediated and represented reality at the detriment of the factual performative engagement with physical reality and the physical bodies of others. This shift might indicate that humans are bound to a more passive and contemplative attitude towards the challenges posed by their relations to the world." [4]

The increase in use of digital media, technologies and applications for communicating seems to be inevitably haunted by the paradoxical duality of being more connected and empowered to communicate than ever but at the apparent price of the quality of our interactions.

This common intuition has been explored in the form of "social media fatigue" [5], predicted depression based on social media use [6] . Most interestingly for this thesis, and specific to the scope of features of digital interactions and what interpersonal dynamics they promote, is the concept of "depersonalizing" [7].

An interesting take on this idea is proposed by Robert Sapolsky in an article where he compares Facebook to Capgras Syndrome [8]. Sapolsky argues that both "factual recognition" and "emotional recognition" are essential to the process of identifying someone as a person we know and if one of the two circuits fail (as it happens in

Capgras Syndrome and as Sapolsky argues happens when faced with a simplistic and approximate representation of another person in Facebook) we are confronted with "cognitive dissonance" and we can't complete the process of recognition to the point we would be able in standard circumstances.

This research tries to propose, and build upon, the hypothesis that an important ingredient for explaning these observed "symptoms" of digital interactions is the lack of "cues" in the virtual representation of others, that in physical life allow us to build an empathetic understanding of who we are interacting with.

1.2.3 Seeing others is understanding others: Mirror Neurons and Empathy

Typing patterns carry important information about what a person is saying but being able to predict and correlate their features with internal state, intentions and meaning through statistical models doesn't inherently mean that these features could be interpreted by another human and serve to enhance the empathy and understanding of the other in a scenario where the interlocutors can see each other's typing. Researches that can support this intuition and motivate us to invest in it further are the ones conducted on "Mirror Neurons". First observed as a set of neurons that fired both when performing and observing the same physical action such as grasping [9], mirror neurons have been linked to similar processes happening with emotions such as disgust [10], suggesting their role in empathy.

"[...] far from being exclusively dependent upon mentalistic/linguistic abilities, the capacity for understanding others as intentional agents is deeply grounded in the relational nature of action." [11]

Perceiving the other performing an action has a fundamental role in understanding them. That's why I think it's motivated to experiment and research new ways of "perceiving" each other in the digital space. Ways that allow us to interpret and empathize with the other.

1.2.4 Possible interventions: Experimental and Critical Design for Well-being

The quote by Vittorio Gallese that opens paragraph B of this state of the art invites us to reflect on what kind of behaviours we're promoting through the "shift" of our habits in the digital era. Most wellbeing apps present on the market today are designed to break existing unhealthy habits [12] without really proposing new paradigms for the user to freely explore, that could lead to the formation of new, persisting healthy behaviours. On the other hand, many brilliant designer and artists have already responded to Gallese's "call to arms for experimental aesthetics", with experimental and critical projects about how we experience the digital world.

Critical designer Ted Hunt has worked for years in projects that try to revolutionize how we see, design and think digital interactions, social media and the internet. "Defaults / The Faults" [13] investigate the role of default suggestions of vehicle for traveling in Google Maps, opening a discussion on the relation between the "defaults" we promote and, for example, climate change.

In "Users Are.. / People Are.." [14] the shift from "person" to "user" is schematized, confronting the features and values that characterize the two categories, highlighting how the role of the user that is perpetuated in most digital applications and services is far from representative of a real human being. The allegedly dangerous simplification denounced in this work reminds of the contrast between "narrow" and "wide" cognitve maps, discussed in Tolman's "Cognitive maps in rats and man" [15].

In the field of text-based communication, it is interesting to point out that, even in the mainstream market, experimental alternatives to traditional models are emerging and gaining significant recognition. "Slowly"[16] is an app where the users chat as if they were distant penfriends, exchanging one "letter" at a time. Another compelling example is "Kialo" [17], a platform designed for rational debates that gives tools (such as tree-like organizations of opinions) specifically tailored for this objective.

Other interesting perspectives on the issue of critical designing for well-being are found in the idea of "embodied values" [18] and "Critical Gameplay" [19].

This thesis aims at prototyping and experimenting with a digital application that follows the principles of these work to try to promote more empathetic interaction in typing-based CMC applications.

Chapter 2

Methods

2.1 Hypothesis and General Methodology

The main research question of this study is: exposure to typing patterns enhance empathic understanding of a partner in a computer mediated text communication.

The experimental design revolves around a one to one online chat-like scenario. The participants are prompted to participate in a remote conversation with a partner. Information about their typing patterns and self-assessed measures of perceived rapport with the partner are collected.

The main objective is to confront the data obtained from a traditional online chat scenario and one in which the users can see each other typing.

The main hypothesis tackles such a broad idea that, in order to focus our efforts and reasonably limit the scope of this research, empathy is operationalized and investigated through these 4 lenses:

- Ability to identify the internal state of the partner
- Perceived interpersonal closeness
- Synchronization of typing patterns
- Revisions and coordination during the dialogue

2.2 Experimental design and set-up

Two custom chat environments have been developed for this experiment. Half of the participants undergo the task interacting through a chat application that mimics the most basic functioning of traditional modern real-time messaging applications. They compose their messages and, once ready, send them. The messages are displayed in a "message bubble" in chronological order. Only text is allowed and no other information about the partner, other than the complete message received, is transmitted.

The remaining participants perform the same task but through another custom made environment. In this one, the messages composed aren't only received once completed and explicitly sent but they appear progressively in real-time, as they're written. Practically, the participants see each other typing the message live (with the latency allowed by the system), revisions (deleting a character) included. The role of the "send" action emitted by a participant (which in the traditional scenario would corresponds to sending the message) in this case will only stop typing in one "message bubble" and create a new one. This will hopefully also contribute to mimic a sense of prosody in the chat and not make it too overwhelming to read as it would be with just a whole uninterrupted stream of text.

Each pair of subjects receives a picture as prompt and are asked to co-author a short story inspired by the content of the image. This prompt is presented in the form of a question asking "What's the story behind this image?" but additional information is provided to explain that, while the picture must be clearly connected to the story, a looser than literal relation with the subject of the picture is allowed. To ensure a consistent measure of the eliciting effect of the picture prompted, the Open Affective Standardized Image Set (OASIS) [20] is used as source image set.

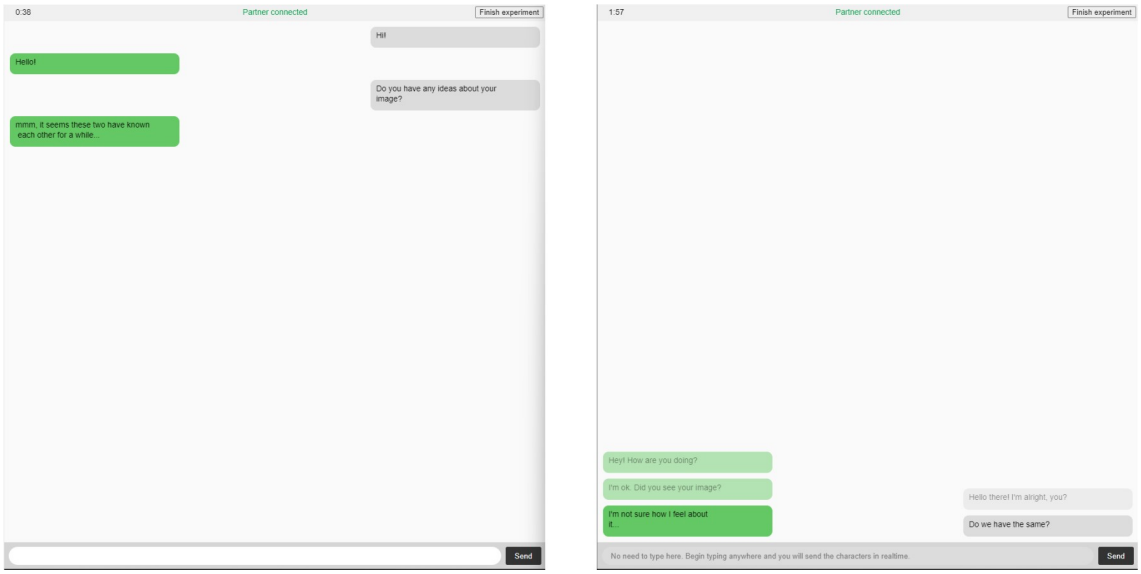

Figure 1: Control group chat (left), Experimental group chat (right)

2.3 Procedures used to obtain data and results

2.3.1 Subjects

A convenience sample of 38 subjects took part to the experiment, 22 females and 16 males, age ranging form 23 to 38, all with similar education levels (mostly bachelor and master). They were divided in 19 couples, 10 assigned to the control group and 9 to the experimental group. The subjects assigned to the same chat didn't know each other and were asked to try and remain completely anonymous by not disclosing any personal detail or information that could give away traits of their identities (only one couple did know each other but they didn't report noticing).

2.3.2 Data collection and analysis

To be able to observe significant modulations in the perceived experience of a participant, a benchmark measure of the subject's personality and empathic tendencies is obtained using the BFI-10, a 10-item Big Five Inventory questionnaire [21] and the Toronto empathy questionnaire [22], a questionnaire of self-reported tendency to feel or understand how the other is feeling in everyday situations. [23]

In both the applications each keystroke has been record with its timestamp. The decision for the features to be extracted and analyzed was based on the previous studies on emotion recognition through keystrokes dynamics presented in the state of the art of this thesis [1] [2]. This data is used to assess synchronization of typing patterns, revisions and coordination during the dialogue. Synchronization analysis servers the purpose of understanding if there's a difference between the two scenarios in the tendency to align with the other typing patterns or another type of correlation between the two patterns (for example one behaviour could consistently elicit another one in the partner). Revisions and coordination (how the two participants take turns, if one is leading, etc.) are analyzed to extract insightful observations about different dyadic dynamics in the conversations.

Once the main task is over, participants are asked to assess their and their partner's arousal and pleasure through the use of the "Affective Slider" [24]. Confronting predicted arousal and pleasure of the partner and their self-assessed counterpart is used to estimate the ability to identify the internal state of the partner.

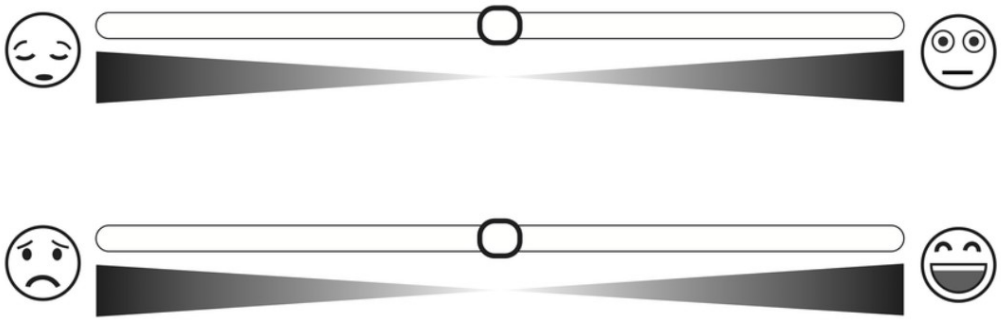

Figure 2: Affective Slider

After this is step is complete, they are asked to complete a graphical questionnaire based on the "Perceived Interpersonal Closeness Scale (PICS)" [25]. This is used to assess any difference in the perceived closeness between the control group and the experimental group. Instead of the discrete categories present in the original PICS questionnaire, we present a gradient representing a continuous value of closeness.

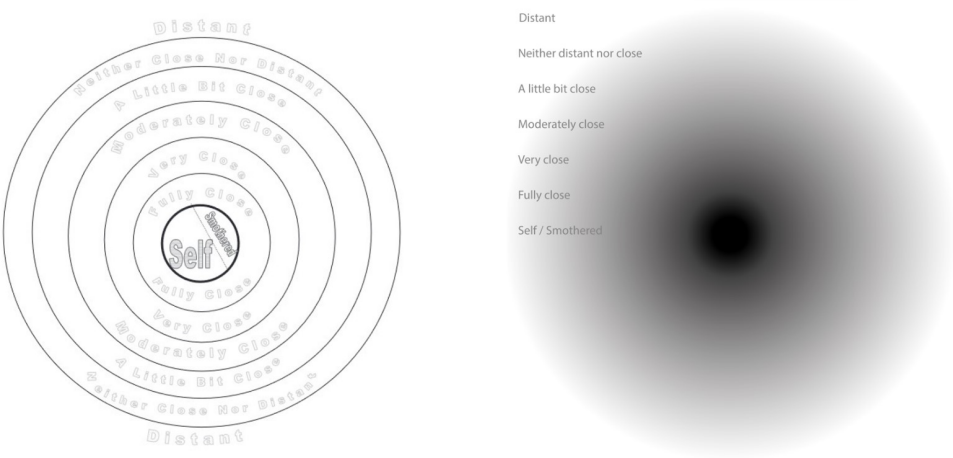

Figure 3: Original PICS (left), Gradient based proposed PICS (right)

Chapter 3

Results

The results are divided in the same categories mentioned in the section 2.1 (Hypothesis and General Methodology)

3.1 Ability to identify internal state of the other

The Affective Slider was presented only to the last 10 couples (5 experimental, 5 control). The results presented in this section are relative to these subjects.

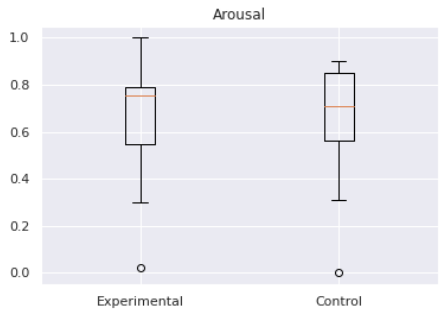

Figure 4: Self-assessed Arousal

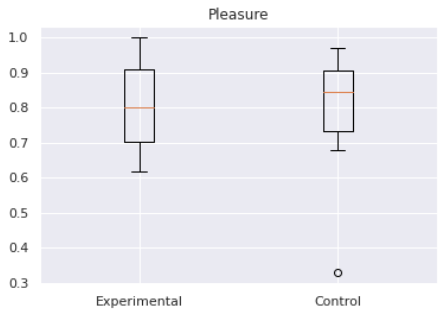

Figure 5: Self-assessed Pleasure

There is no significant difference between the experimental and control groups in the self-assessed levels of arousal and pleasure reported through the Affective Slider. After testing for normality, an independent samples t-test reports for arousal *(statistic=0.16, pvalue=0.87)* and for pleasure *(statistic=-0.55, pvalue=0.59)*.

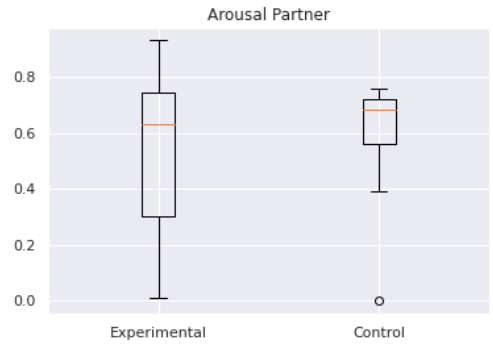

Figure 6: Predicted partner's Arousal

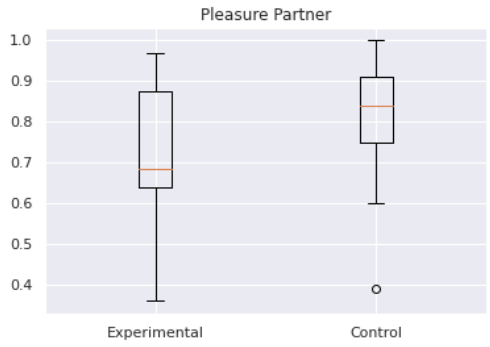

Figure 7: Predicted partner's Pleasure

For the predicted arousal of the partner a significant difference in variance is found between experimental and control group. After testing for normality, a Bartlett test for equal variances reports *(statistic=5.94, pvalue=0.02)*. Running the same test for predicted partner's pleasure returns a similar result of *(statistic=5.38, pvalue=0.02)*. For predicted pleasure, also the mean is significantly different between the two groups *(statistic=6.22, pvalue=0.01)* (result of a AlexanderGovern test for equal means, for samples with heterogeneity of variance.)

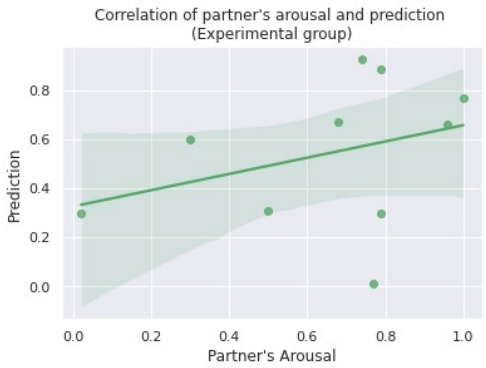

Figure 8: Predicted/Actual arousal correlation experimental group

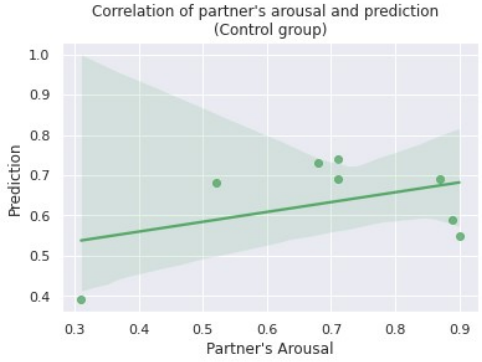

Figure 9: Predicted/Actual arousal correlation control group

While plotting correlations between predicted and actual values of arousal and pleasure seem to suggest positive correlation, none of these prove statistically significant after running Spearman tests. Fig. 8: *(correlation=0.32, pvalue=0.36)*, Fig. 9: *(correlation=-0.03, pvalue=0.94)*, Fig. 10: *(correlation=0.01, pvalue=0.97)*, Fig. 11: *(correlation=0.30, pvalue=0.43)*.

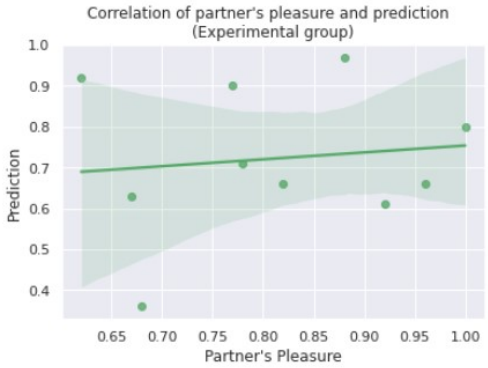

Figure 10: Predicted/Actual pleasure correlation experimental group

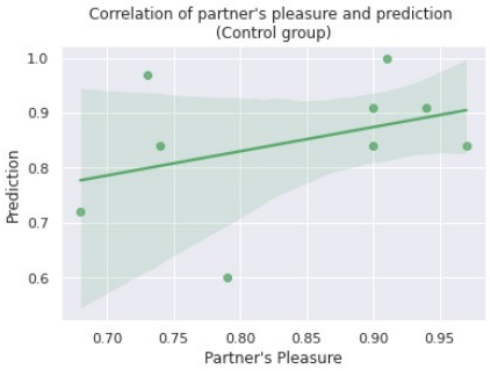

Figure 11: Predicted/Actual pleasure correlation control group

3.2 Perceived interpersonal closeness

Most participants reported having troubles understanding and filling the gradient variation of the PICS questionnaire. Some delivered it with significant delay and others didn't at all. Nonetheless, the average distance of the partner from the center of the gradient has been measured and there's no significant difference between the two groups. *(statistic=0.12, pvalue=0.73)*

3.3 Revisions and coordination during the dialogue

While average keystrokes per minute don't differ significantly between the two groups *(statistic=-0.78, pvalue=0.44)*, the null hypothesis for the equality of variances can be rejected when confronting the distributions of difference in keystrokes per minutes between partners: Bartlett test reports *(statistic=4.40, pvalue=0.04)*.

We consider a revision initiated when it's the first time the "Backspace" key is pressed in a uninterrupted sequence of "Backspace" presses. A difference close to significance is observed between the two groups when analyzing how many times they "initiated a revision" when testing it with a One-way ANOVA *(statistic=3.04, pvalue=0.09)*, after testing for normality and homogeneity of variances.

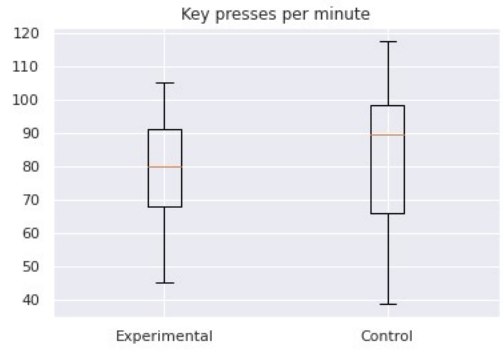

Figure 12: Average keypresses per minute

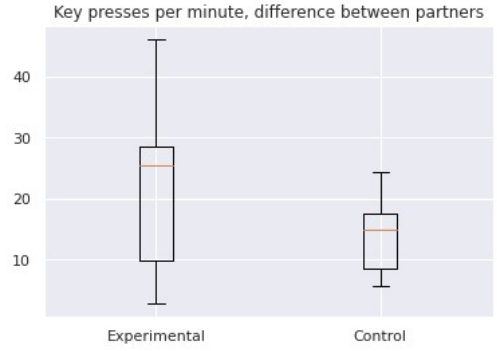

Figure 13: Difference between partners of average keypresses per minute

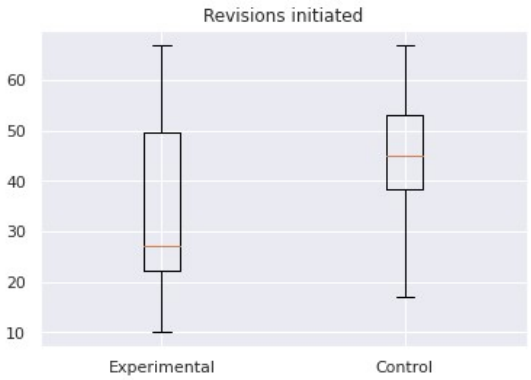

Figure 14: Revisions initiated

3.4 Synchronization of typing patterns

Logging every keystroke sequentially from the chat proved to be a very naive way of collecting data for the scope of this experiment. This part of the analysis required a more complex and high level representation of the conversation in terms of turn taking. No results regarding synchronization could be confidently extracted.

Chapter 4

Discussion

The problem statement of this thesis presented an incredibly broad panorama of ideas and perspectives worth confronting in the area of digital human to human communication. A big effort in this research has been dedicated to trying to reconcile the many different focal points of the discussion, in an attempt to operate with an holistic approach.

Building up the experiment from a relatively basic scenario but with the consolidated background of typing patterns research, started tracing interesting path that deserve more attention.

4.1 Discussion

Even though the results made us reject some of our higher level hypothesis (such as expected significant effects on the understanding of the other and the perceived interpersonal closeness), digging deeper in the results reveal some possibly interesting implications.

The significant difference in variances of the predicted arousal and pleasure data between experimental and control group (figure 6 and 7), suggest a broader range of predicted values proposed by subjects in the experimental group, while the control group seems to pick from a more limited spectrum. This strategy doesn't signifi-

cantly favour their accuracy as we see in the correlation analysis.

A similar result is noticed when analyzing the difference of average key-presses per minutes between partners. While still tending towards a cooperative/balanced rhythm, the experimental group tend to reach higher differences in key-presses per minute, suggesting that subjects in the experimental group were more prone to take the lead / let the other take the lead in the discussion.

Lastly, while only close to significance (p=0.09), the difference in number of initiated revisions between the two groups could indicate a more spontaneous attitude towards the conversation.

4.2 Conclusions

One big limitation to this study was the inability to process smaller units of the dialogue (such as turns) in the conversation. While an overall analysis of the whole discussion already pointed in some interesting directions, a closer attention to internal dynamics in the dialogue is required to really understand how the discussion evolves and mutates through the course of the experiment.

Another step that could push this research further would be analyzing semantic elements of the conversation and analyze what type of correlation they have with the dynamics that have been studied in this thesis.

Nonetheless, it's interesting to see significant results appear as relative "widening" and "narrowing" of the ranges of values analyzed, rather than in a complete shift of the distribution. It's tempting to fantasize about possible connections with the idea of Cognitive and Emotional Maps elaborated by Edward C. Tolman and the almost "elastic" behaviour they present in certain environment or situations.

The still primitive and inconclusive hints that we get from these results are not enough to firmly reject the null hypothesis of this thesis but support a minor implicit hypothesis: we're not unaffected by the digital mean that hosts our interactions. Language and empathy, as we've seen in the state of the art of this thesis,

are complex phenomenons, far from the rationalist view of the purely logical deconstruction of the content of a message. Language is a sensual, physical and embodied event. If we don't want to regress into unaffective forms of communication we need to invest in questioning our digital environments and constantly strive for a deeper and profound connection with the other.

List of Figures

1 Control group chat (left), Experimental group chat (right) 9
2 Affective Slider 10
3 Original PICS (left), Gradient based proposed PICS (right) 11
4 Self-assessed Arousal 12
5 Self-assessed Pleasure 12
6 Predicted partner's Arousal 13
7 Predicted partner's Pleasure 13
8 Predicted/Actual arousal correlation experimental group 13
9 Predicted/Actual arousal correlation control group 13
10 Predicted/Actual pleasure correlation experimental group 14
11 Predicted/Actual pleasure correlation control group 14
12 Average keypresses per minute 15
13 Difference between partners of average keypresses per minute 15
14 Revisions initiated 15

Bibliography

[1] Yang, L. & Qin, S.-f. A review of emotion recognition methods from keystroke, mouse, and touchscreen dynamics. *IEEE Access* (2021).

[2] López-Carral, H., Santos-Pata, D., Zucca, R. & Verschure, P. F. How you type is what you type: Keystroke dynamics correlate with affective content. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–5 (IEEE, 2019).

[3] Goodkind, A. Using keystrokes to predict social dynamics in dialogue (2021).

[4] Gallese, V. The aesthetic world in the digital era. a call to arms for experimental aesthetics. *Reti, saperi, linguaggi* **7**, 55–84 (2020).

[5] Dhir, A., Kaur, P., Chen, S. & Pallesen, S. Antecedents and consequences of social media fatigue. *International Journal of Information Management* **48**, 193–202 (2019).

[6] De Choudhury, M., Gamon, M., Counts, S. & Horvitz, E. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media* (2013).

[7] Ciaunica, A. *et al.* Zoomed out? depersonalization is related to increased digital media use during the covid-19 pandemic lockdown (2021).

[8] Sapolsky, R. To understand facebook, study capgras syndrome (2016). URL `https://nautil.us/to-understand-facebook-study-capgras-syndrome-5374/`.

[9] Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996).

[10] Wicker, B. *et al.* Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron* **40**, 655–664 (2003).

[11] Gallese, V. The shared manifold hypothesis. from mirror neurons to empathy. *Journal of consciousness studies* **8**, 33–50 (2001).

[12] Monge Roffarello, A. & De Russis, L. The race towards digital wellbeing: Issues and opportunities. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–14 (2019).

[13] Hunt, T. Defaults / the faults (2020). URL `http://ted-hunt.com/DEFAULTS.html`.

[14] Hunt, T. Users are.. / people are.. (2018). URL `http://ted-hunt.com/USERS-PEOPLE.html`.

[15] Tolman, E. C. Cognitive maps in rats and men. *Psychological review* **55** (1948).

[16] Ltd., S. C. "slowly" (2018). URL `https://slowly.app/`.

[17] Kialo Inc. "kialo" (2018). URL `https://kialo.com`.

[18] Flanagan, M., Howe, D. C. & Nissenbaum, H. Embodying values in technology: Theory and practice. *Information technology and moral philosophy* **322**, 24 (2008).

[19] Grace, L. D. Critical gameplay: designing games to critique convention. In *Proceedings of the 20th ACM international conference on Multimedia*, 1185–1188 (2012).

[20] Kurdi, B., Lozano, S. & Banaji, M. R. Introducing the open affective standardized image set (oasis). *Behavior research methods* **49**, 457–470 (2017).

[21] Rammstedt, B. & John, O. P. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of research in Personality* **41**, 203–212 (2007).

[22] Spreng*, R. N., McKinnon*, M. C., Mar, R. A. & Levine, B. The toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment* **91**, 62–71 (2009).

[23] Surma-Aho, A. & Hölttä-Otto, K. Conceptualization and operationalization of empathy in design research. *Design Studies* **78**, 101075 (2022).

[24] Betella, A. & Verschure, P. F. The affective slider: A digital self-assessment scale for the measurement of human emotions. *PloS one* **11**, e0148037 (2016).

[25] Popovic, M., Milne, D. & Barrett, P. The scale of perceived interpersonal closeness (pics). *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice* **10**, 286–301 (2003).