Watching a TV Program Together with a Partner in a Remote Location

Matej Kaninsky

Part 1

INTRODUCTION

Although watching a traditional TV is in decline among younger audiences, the rise of video-on-demand services (e.g. Netflix, Amazon Prime, YouTube) [22] suggests that watching video programmes is still an important part of their leisure activities.

Sharing the experience of watching TV/video through the use of video-mediated communication was found to be one of the ways of staying connected in long-distance relationships with partners, friends or family [4].

Video-mediated communication (VMC) plays a crucial role in long-distance relationships in general [20] and given the prevalence of watching TV and video programs [22], it is surprising how little research has been done to understand needs and barriers that partners face when watching video programs remotely together.

This study aims to build on Macaranas *et al.'s* research [15] that was among first to systematically examine various remote co-watching settings. It did not, however, address thoroughly the affective states participants experienced.

An experiment followed by semi-structured interviews and video emotion-analysis was conducted with a married couple, that often travels separately.

RELATED WORK

Macaranas *et al.* [15] conducted a study on how people use VMC to watch video programs together. They first ran a field study with 56 participants (29 male, 27 female), in which the participants engaged in watching a TV in pairs while being on Skype. Second study they conducted was a within-subject lab experiment with eight pairs of participants (6 male, 10 female) to explore the differences of watching TV together in three conditions: (1) being in the same room (*local condition*), (2) being in different rooms and using one device for VMC and a video (*picture-in-picture (PIP) condition*), (3) being in different rooms using different devices for VMC and a video (*proxy condition*).

In the field study, most participants (61%) chose a *PIP* configuration over 39% who picked the *proxy*. A statistical analysis of post-experiment survey did not reveal any significant differences in enjoyment. In the lab study, however, participants preferred the *proxy* and *local* setting over *PIP*.

The study provided valuable findings in a new field of research but focused primarily on usability issues and exploration of viewing settings; this presents an opportunity for expanding the study by examining the affective layer of the experience.

Two research questions (RQ) were formed to help structure the coursework study:

RQ1: What affective states do people experience while interacting with each other when watching video together remotely?

RQ2: How do they respond to those affective states?

METHOD

Pilot testing

Using two separate rooms, I conducted a pilot test with my acquaintance to assess the feasibility of *PIP* and *proxy* viewing settings. The *proxy* setting suffered from a strong audio cross-talk (echoes), which would make it impossible for the participants to communicate naturally. The *PIP* setting was therefore selected for the main experiment.

Several online services offering synchronised remote video watching were also evaluated during the pilot – *rabb.it*, *netflixparty.com*, *showgoers.tv* and *letsgaze.com*. Only *letsgaze.com* offered a picture-in-picture functionality and was therefore chosen for the experiment.

Participants

A married couple (female, 28 years – P1; male 31 years – P2) was selected to participate in the study through an on-line screening questionnaire (see Appendix 1) which was distributed via social media to my acquaintances. Fifteen responses were obtained and the following criteria for selection were used: (1) has used a mainstream VMC in the past month, (2) finds watching TV/video programs with other people enjoyable, (3) has not watched a video program with another person remotely before but would be interested in doing so.
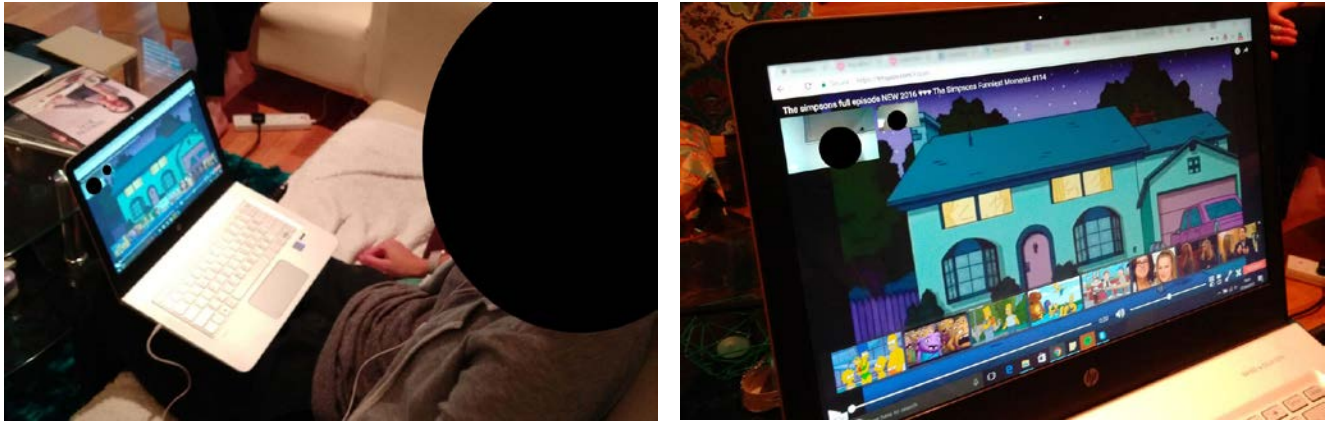
Figure 1. Left: Experiment setting; Right: *PIP* condition

Apparatus

Two Windows 10, 15.6" laptops were used, running a *letsgaze.com* service in a Google Chrome browser.

Two video streams of both participants were recorded for further analysis: (1) a laptop screen; (2) facial expressions of each participant.

Procedure

Participants were situated in two separate flats in London. Sitting on a sofa in a living room (see Figure 1), they were briefed and asked to sign an informed consent (see Appendix 2). Then, they were given time to select a video they would both want to watch – they selected one episode of The Simpsons.

Next, they spent approximately 20 minutes watching, being each alone in the room. The participants were briefed to behave naturally and told they would not be tested on the content of the video.

Right after the experience, a semi-structured interview was carried out with each participant in turn.

Semi-structured interviews

In order to understand affective states experienced during the experiment, several interview techniques and frameworks were embedded into the interview process and questions.

Mauss and Robinson [16] in their review article on measures of emotion suggest that self-reports of emotion are generally more valid when applied to the currently experienced emotions or directly after an event. Interviews were therefore conducted right after the experiment.

The interview structure followed the *Spatio-temporal thread* of the McCarthy and Wright's *Technology as experience* framework [17] – the participants were guided to relive the experience from the beginning to the end. This was combined with Petitmengin's interview method [24] which

aims to elicit precise descriptions of participants' subjective experiences. Following this method, participants' attention was first stabilised by setting expectations about the length of the interview and its focus (the interaction between the participants, not the video content). Then, they were guided to reconstruct and describe the interactions between them during the experiment.

Building on the related research of social presence [1,27] and ambient social TV [11], measures of social presence and co-presence from the *Networked Minds Social Presence Inventory* [2] which normally serve as questionnaire scales, were adapted and incorporated into the interview to explore the sense of co-presence between participants.

Lastly, at the end of the interview, the participants were asked to reflect on the experience as a whole following Norman's three-tier design framework which includes *visceral*, *behavioural* and *reflective* levels [21]. Norman argues that only the *reflective* level requires conscious interpretation. It is, therefore, suitable to be explored in an interview.

See a discussion guide in Appendix 3.

RESULTS & DISCUSSION

The interviews were transcribed verbatim using NVivo 11 and analysed following the thematic analysis methodology [3] and Grounded theory methodologies of open coding, memoing and mind-mapping [6] (see Appendices 4, 5, 6, 7). The participants did not find the experience very enjoyable, which reflects in the six overarching themes identified:

1. Privacy intrusion
2. Situational objective self-awareness
3. Boredom
4. Divided attention
5. Obligation to interact
6. Expecting emotion reciprocity

Each theme was examined and affective states were extrapolated based on literature review findings, participants' self-report and analysis of the videos from the experiment. Identified affective states were confirmed with the participants in a follow-up session in which a video of the experiment was re-played.

Before assigning specific names to identified affective states, three seminal affective state taxonomies were reviewed: (1) *Plutchik's three-dimensional model* [25,26], (2) Rusell and Barrett's *Core affect model* [28], and (3) Ortony, Clore and Foss's *Affective Lexicon* [23].

Below, three of the six themes are discussed. For reference, please visit Appendix 8 to see the discussion of the other three themes.

1. Privacy intrusion: Anxiety, Vigilance, Alertness
Both participants were aware of being watched by each other, which had an impact on their behaviour and interactions inducing mild situational anxiety, vigilance and alertness. One of the social anxiety symptoms is feeling uncomfortable while being watched [19]:

> I'm kind of being watched while watching a TV...[]...[When sitting next to each other] you don't have that direct eye contact on you. (P1)

In a large-scale study [10], researchers compared data sets of 63 countries and among other factors looked at correlations between autonomy and well-being. They found that especially in Western countries where people are more individualistic decreased autonomy led to increased feelings of anxiety. Similar results were found in workplaces where surveillance technologies were used [14].

In this experiment, P2 felt more observed than he was observing and described decreased autonomy:

> Because I knew she was looking at me, so I had to focus, I couldn't be doing another thing.... (P2)

> []...I felt that she was looking more at me because she was saying things like... when I was taking my phone and I was only moving my eyes, so I felt observed a little bit. More than I was observing. (P2)

2. Situational objective self-awareness: Discouragement, Anxiety, Vigilance
In a recent study [12], researchers showed that seeing oneself in a video-chat led to increased cognitive load and had negative consequences for communication and behaviour in professional teams. Research on self-awareness and especially the Objective self-awareness theory [8] has shown that when people observe themselves in a mirror (as well as on a monitor), they tend to perceive themselves as an object (objective self-awareness), which may impact performance similarly to test anxiety [12].

P2 was affected by objective self-awareness (OSA) and that had a negative impact on his experience – vigilance, anxiety:

> I was more aware of seeing myself there and I was probably not smiling so much and not laughing so much as I would do if I didn't see my face. It was kind of a... kind of distracting me. (P2)

P2, who felt more self-aware than P1 put in place emotion-regulation strategies to decrease the distracting effects of OSA - he felt discouraged to show his facial expressions and body movement and tried to control them:

> And I also thought I would distract the other person if I would... change my expressions a lot or... It made me stay more still than I would normally be. I was smiling a little bit and then I was seeing myself smiling and then I was controlling myself. (P2)

3. Boredom
Although both participants agreed on the video to watch, P2 was *"not as excited about the video as her"*, which made him feel bored at times, so he started looking for other sources of stimulation – for example switching to using his smartphone. This was not received positively by P1, however:

> []...and then he started reaching for his phone which always annoys me because then he always asking me later what's going on, so I told him not to look at his phone [laughing]. (P2)

Boredom will be described in more detail in *Part 2* as well as anxiety.

4. Divided attention: Overwhelmedness
In contrast with Macaranas *et al.*'s study [15], which suggested there was little evidence for divided attention between video watching and Skype communication, participants in this study expressed overwhelmedness caused by focusing on the two different channels at the same time:

> []...you kind of concentrate on watching but you kind of feel like you have to somehow interact a bit more with the person because you're still having like a Skype conversation. (P1)

> []...if were to do this every day, I don't know, I'm assuming now, it may be a little bit more overwhelming or tiring to feel like this [being watched] all the time. (P2)

P2 tried to overcome this issue by focusing more on the video only (but that was not received well by P1 who expected emotion reciprocity (theme #6).

5. Obligation to interact: Encouragement (negative valence)
Seeing her partner in the video all the time encouraged P1 initiate more interactions – verbal and non-verbal – than she

would do in a local setting. This was not perceived in a positive way, however:

> *[]... I personally felt more obliged... to have a bit of more conversation. Otherwise it would be just a bit weird, watching like a Skype window and a TV...[]... I felt like I had to be more proactive in the interaction. (P1)*

6. Expecting emotion reciprocity: Disappointment, Anxiety

Obligation to interact connects tightly to emotion reciprocity. Buunk and Schafeli [5] argue that *"humans have developed innate mechanisms to expect reciprocity in interpersonal relationships and that a lack of reciprocity is accompanied by negative affect"*. P1 who initiated most interactions felt disappointed from the lack of reactions back.

> *[]...you think something is funny and the other person does not, so I think even in the real interaction, that will be a bit like - 'is it me, am I weird?' (P1)*

> *Yeah... I think I was trying, I was a bit like... by making comments about the video or something, like 'we haven't watched this, this is new'... So yeah, initiating a bit more of like some reactions from his side. And I mean he was communicating with me but he wasn't like laughing or anything. (P1)*

TECHNOLOGY EVALUATION: LETSGAZE.COM

In this section, *Letsgaze.com*, a synchronised video watching service, that was used as the main technology during the experiment, is evaluated in the light of identified affective states.

The service provides a very simple interface and the only setting that viewers can change is toggling between seeing faces of both partners or seeing neither.

Therefore, the discussion about how the interface can impact the identified affective states is reduced to the impact of seeing or not seeing oneself and each other.

Not using the video chat could ease the anxiety, vigilance and alertness caused by *being watched*. However, when asked, the participants did not prefer that option since it interfered with their desire to spend time together.

Switching the video chat off could reduce the feelings of objective self-awareness, overwhelmedness and disappointment from the lack of interaction as well but it is clearly a sub-optimal solution since it hampers the experience of mutual closeness and connectedness.
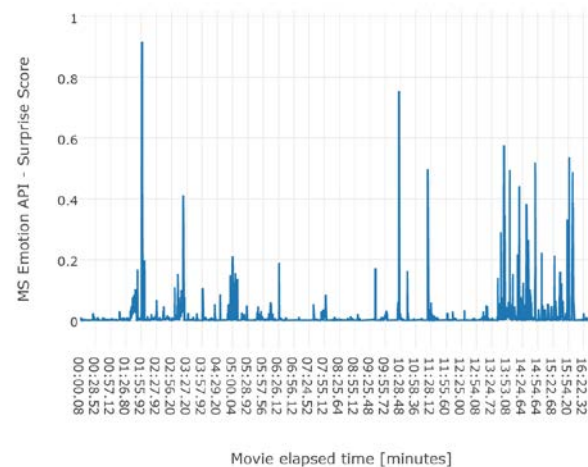
Boredom is not addressed by *Letsgace.com* at all.

Figure 2. MS Emotion API – False detections of surprise

FACIAL EXPRESSION ANALYSIS

MS Emotion API[1] was used to analyse facial expressions of the participants captured during the experiment.

The API returns eight basic emotion values. Although sounding promising the results were disappointing. When manually comparing the videos and detected emotions, there were obvious glitches (e.g. returning *surprise* when a participant touched her mouth – see Figure 2). The data overall did not correlate well with manual assessment and is not therefore reported. Although a valuable lesson was learnt – the technology is still not ready to supplement research.

CONCLUSION

Eight affective states were identified as an answer to RQ1: alertness, anxiety, boredom, disappointment, discouragement, encouragement (negative valence), overwhelmedness and vigilance. Although the participants also experienced emotions with positive valence, these were not taken into consideration because they were identified as being connected to the content of the video itself and not the interpersonal interaction which is of primary concern.

To answer RQ2, the participants put in various strategies which often went against each other and further made the experience more unsatisfactory, e.g. initiating interactions (P1) vs staying still (P2).

Limitations of the study

Small sample
This study only reflects findings from two participants. That has an impact on ecological validity. Best efforts were

---

[1]https://www.microsoft.com/cognitive-services/en-us/emotion-api

employed, however, to ground the findings and conclusions in existing literature.

One researcher
Since coursework projects are individual, I was the only researcher. In a study like this, however, multiple coders should review the transcripts to counterbalance each other's biases.

Physical measures
I also considered using measures of GSR and heart rate to further inform the study. These could be explored to validate whether there are any physiological patterns when participants verbally or non-verbally interact.

Part 2

INTRODUCTION
For further analysis, two affective states are considered: (1) **situational social anxiety** in relation to autonomy, and (2) **boredom**. These two states had a big impact on how the participants were emotionally experiencing the experiment. Both states have negative valence and therefore there is an opportunity for a technology intervention to reduce or ideally remove these affective states from the experience.

SITUATIONAL SOCIAL ANXIETY
Both participants showed symptoms of situational social anxiety (none of the participants is suffering from clinical anxiety). These were not overwhelming for them but caused discomfort and a change of behaviour. Among these symptoms were[2]:

- Finding it difficult to behave naturally while being watched
- Feeling of being judged while being watched
- Avoiding eye contact

Self-Determination Theory and especially its sub-theory *Cognitive Evaluation Theory* (CET) [29] suggest there are three factors to person's intrinsic motivation: (1) competence, (2) autonomy, and (3) relatedness. Intrinsic motivation is key to people's engagement and enjoyment of an activity. At least one of the factors above was largely violated in the experiment – there were clear constraints imposed on participants' autonomy by being able to observe each other – as P2 noted: "We were kind of controlling the other".

The effect on overall experience
Low levels of anxiety were part of two identified themes: (1) Privacy intrusion, and (2) Situational objective self-awareness. Experiencing this affective state is interesting especially because the participants are married and are used to be around each other. We could, therefore, infer that anxiety was situational and can be ascribed to the experimental condition.

When reflecting on the experience as a whole, both participants expressed that they either would not want to see themselves all the time (P2) or would like to introduce some internal rules to set expectations for the experience (P1). These are certainly not signs of a pleasant experience.

Design implications
Taking into consideration the technology used in the main experiment, *Letsgaze.com*, a simple recommendation could be to offer the users control over the viewing settings, i.e. they could select from seeing: (1) self and partner, (2) partner only, (3) no VMC. This, however, may have a negative impact on the closeness of participants during the experience, which is the main reason why participants engage in this activity in the first place [15].

If participants would not see each other during the experience, could they experience the sense of closeness in a different way? Ambient media, e.g. light and sound [13], have been used to support ambient awareness as part of tangible user interfaces (TUI) [30]. For instance, Ambient social TV [11] enhances a traditional TV by using an ambient lamp that sits next to the TV and by a change of colours communicates if user's friends are also currently watching.

An ambient telephone project [9] explored the use of light intensity to communicate arousal of the communication, and different colours to identify callers.

Two streams of design could be explored to support ambient awareness:

(1) Tangible user interfaces

While watching a synchronised video stream on a TV or other device, users could use an external device that could use light intensity and colour as means to communicate presence to each other.

Another option to explore could be a tangible device that would warm up to communicate presence, this could further enhance a sense of intimacy that long-distance relationships lack.

---

[2] http://www.nhs.uk/conditions/social-anxiety/Pages/Social-anxiety.aspx

(2) Graphical user interfaces

Staying within a more traditional web and app design, possible directions worth exploring include adding a coloured strip or dot to ambiently represent the other user, replacing the video chat. Presence could be communicated by changing colour hue and intensity.

If a device had a touch display, a simple touch on the strip/dot could communicate presence and closeness to the other user.

BOREDOM

Mikulas and Vodanovich [18] define boredom as "a state of relatively low arousal and dissatisfaction, which is attributed to an inadequately stimulating situation".

If bored, people will seek optimal levels of arousal and will switch tasks to achieve that optimal state [18].

Csikszentmihalyi [7] argues that in boredom "the low level of challenge relative to skills allows attention to drift".

Symptoms of boredom (switching to a more stimulating task) were observed on P2's behaviour during the experiment.

The effect on overall experience

After a few minutes into the video, P2 started looking at his phone and repeated that several times during the experience. The participant did not find the video stimulating: *"I was not as excited about the video as her."* He would, however, continue watching to spend time with his partner.

A negative consequence of this behaviour arose by P1 seeing him using the phone and telling him not to, which left P2 with no other option (should he not want to engage in a conflict) than trying to focus on the video watching activity which he did not find stimulating.

Design implications

The intermittent boredom of one of the users may be inevitable during the course of a long video. When sharing the same physical location, a bored user can target their attention to other stimuli without causing misunderstanding, in a video-mediated communication, however, misinterpretation of user's intentions is more common [4]. This can then easily lead to a conflict as shown in the presented study.

To engage a bored user, designers should consider offering activities alternative, more stimulating, activities, while keeping them unobtrusive to the other user.

To help the user enter a *flow* state, several conditions have to be met [7]:

- Clear and immediate feedback
- Match between perceived skills and perceived challenges

- Sense of control
- Clear set of goals

Several design scenarios could be considered:

A set of activities related to the program watched could be available to the user should they get bored. These could have for example a form of interesting facts about the program, that could be displayed as snap messages on the screen.

Since reading takes more attentional resources [31] than listening, the service could provide an option to switch languages and displaying subtitles, so that users could independently watch the program in different languages. This may allow a person who gets bored to practise a second language, which may be more stimulating.

CONCLUSION

Two affective states were discussed and options for future technology developments presented. Although grounded in literature, the design directions would need to be validated with users to assess their benefits.

Part 3

REFERENCES

[1]   Biocca, F. and Harms, C. Defining and Measuring Social Presence - Contribution to the Networked Minds Theory and Measure. *International Workshop on Presence*, 517 (2002), 1–36.

[2]   Biocca, F. and Harms, C. Networked Minds Social Presence Inventory (Scales only). (2003).

[3]   Braun, V. and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (2006), 77–101.

[4]   Brubaker, J.R., Venolia, G., and Tang, J.C. Focusing on shared experiences. *Proceedings of the Designing Interactive Systems Conference on - DIS '12*, ACM Press (2012), 96.

[5]   Buunk, B.P. and Schaufeli, W.B. Reciprocity in Interpersonal Relationships: An Evolutionary Perspective on Its Importance for Health and Well-being. *European Review of Social Psychology 10*, 1 (1999), 259–291.

[6]   Charmaz, K. *Constructing grounded theory*. Sage, 2014.

[7]   Csikszentmihalyi, M. *Flow and the Foundations of Positive Psychology*. Springer, 2014.

[8]   Duval, S. and Wicklund, R.A. A theory of objective self awareness. (1972).

[9]   Emparanza, J.P., Dadlani, P., De Ruyter, B., and

Härmä, A. Ambient telephony: Designing a communication system for enhancing social presence in home mediated communication. *Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, (2009).

[10] Fischer, R. and Boer, D. What is more important for national well-being: money or autonomy? A meta-analysis of well-being, burnout, and anxiety across 63 societies. *Journal of personality and social psychology 101*, 1 (2011), 164–184.

[11] Harboe, G., Metcalf, C.J., Bentley, F., Tullio, J., Massey, N., and Romano, G. Ambient Social TV: Drawing People into a Shared Experience. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, ACM Press (2008), 1.

[12] Hassell, M.D. and Cotton, J.L. Some things are better left unseen: Toward more effective communication and team performance in video-mediated interactions. *Computers in Human Behavior 73*, (2017), 200–208.

[13] Ishii, H. and Ullmer, B. Tangible bits. *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 97 39*, (1997), 234–241.

[14] Kizza, J.M., Ssanyu, J., and Danielson, P. Electronic Monitoring in the Workplace: Controversies and Solutions. 2005.

[15] Macaranas, A., Venolia, G., Inkpen, K., and Tang, J. Sharing Experiences over Video: Watching Video Programs together at a Distance. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2013, 73–90.

[16] Mauss, I.B. and Robinson, M.D. Measures of emotion: A review. *Cognition and Emotion 23*, 2 (2009), 209–237.

[17] McCarthy, J. and Wright, P. Technology as experience. *interactions 11*, 5 (2004), 42.

[18] Mikulas, W.L. and Vodanovich, S.J. The essence of boredom. *The Psychological Record 43*, 1 (1993), 3.

[19] National Health Service (NHS). Social anxiety (social phobia). 2017. http://www.nhs.uk/conditions/social-anxiety/Pages/Social-anxiety.aspx.

[20] Neustaedter, C. and Greenberg, S. Intimacy in long-distance relationships over video chat. *Proceedings of the SIGCHI Conference on …*, (2012), 753–762.

[21] Norman, D.A. *Emotional design: Why we love (or hate) everyday things*. Basic Civitas Books, 2004.

[22] OFCOM. *PSB Annual Research Report 2016*. 2016.

[23] Ortony, A., Clore, G.L., and Foss, M.A. The Referential Struture of the Affective Lexicon. *Cognitive Science 11*, (1987), 341–364.

[24] Petitmengin, C. Describing one's subjective experience in the second person : An interview method for the science of consciousness. (2006), 229–269.

[25] Plutchik, R. A general psychoevolutionary theory of emotion. *Theories of emotion 1*, 3–31 (1980), 4.

[26] Plutchik, R. The nature of emotions: Human emotions have deep evolutionary roots. *American Scientist 89*, 2001, 344–350.

[27] Rae, I., Venolia, G., Tang, J.C., and Molnar, D. A Framework for Understanding and Designing Telepresence. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, ACM Press (2015), 1552–1566.

[28] Russell, J.A. and Barrett, L.F. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology 76*, 5 (1999), 805–819.

[29] Ryan, R. and Deci, E. Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist 55*, 1 (2000), 68–78.

[30] Shaer O.a Hornecker, E.. Tangible User Interfaces: Past, present, and future directions. *Foundations and Trends in Human-Computer Interaction 3*, 1–2 (2009), 1–137.

[31] Tindle, R. and Longstaf, M.G. Writing, Reading, and Listening Differentially Overload Working Memory Performance Across the Serial Position Curve. *Advances in Cognitive Psychology 11*, 4 (2015), 147–155.

APPENDICES (Excluded from the online portfolio version)