The Effect of Task and Ownership on Time Estimation in Virtual Environments

Karin Hägni[1,2], Kynan Eng[1], Marie-Claude Hepp-Reymond[1],
Lisa Holper[1], Birgit Keisker[3], Ewa Siekierka[3] and Daniel C. Kiper[1]

[1] Institute of Neuroinformatics, University/ETH Zurich, Switzerland
[2] Institute of Human Movement Science and Sport, ETH Zurich, Switzerland
[3] Department of Neurology, University Hospital Zurich, Switzerland

Abstract

To date, efforts to quantify presence using subjective time estimation have met with mixed results. In this paper we present experiments investigating the relation between reported levels of presence and prospective time estimation, defined as self-termination of a given task after a perceived elapsed time. We conducted a study with 58 healthy subjects performing four motor tasks with putatively different levels of presence. Our results show that more engaging tasks yield higher reported levels of presence and lead to slower perceived passing of time (i.e. longer elapsed time with respect to actual time passed). Tasks which involve higher levels of ownership also elicit longer real elapsed times. Prospective time estimation is affected by the degree to which a task is goal-oriented, and by the order in which successive tasks are performed. Our results support the use of prospective time estimation as a measure of presence and/or task involvement.

Keywords--- **Presence, immersion, prospective time estimation, involvement, ownership**

1. Introduction

Many different models have been introduced to account for the cognitive response of a subject *immersed* in a virtual environment, often referred to as the subject's *presence* in the environment. We take Slater's [1] proposal as our working definition of immersion as being the total input stimulation to a subject provided by a particular set of technologies, and presence being the overall effect on a subject of receiving input of a particular immersion. In this definition, presence is distinct from *involvement*, which is related to the content of the virtual environment. For example, it is possible to be highly present in a virtual environment equipped with highly immersive technologies, but if the content is not interesting then one's involvement will be correspondingly low. We also define *ownership* as a particular type of presence in which the participant perceives part of the immersion as being under his/her control, e.g. virtual limbs which respond realistically to the user's limb movements.

Models of presence can be grouped into those which assume a priori categories or dimensions, and those which try to at least partly define those categories on the basis of measurements. One example of the former is a model proposed by Waterworth and Waterworth [2], which employs a three-dimensional space representing a person's momentary state of mind. The three dimensions are i) locus, the degree to which the subject focuses on the real or the virtual world, ii) focus, the extent to which the subject is focused on external (real or virtual) stimuli or internal processes, and iii) sensus, which takes the subject's state of arousal into account (from unconscious to alert). Another model proposed by Sas and O'Hare [3] focuses primarily on the end user's personality traits rather than on the stimuli presented by an immersive system. They define four factors underlying the subjective experience of presence: 'absorption', 'creative imagination', 'empathy' and 'cognitive style'.

Presence models which attempt to define categories on the basis of measurements typically use questionnaires. This method simultaneously takes the subject's personality and the system's immersive capabilities into account. It allows the structuring of the various aspects of the experience of presence into categories. Interestingly, Schubert et al. [4] and Lessiter et al. [5] have found very similar categories. Schubert et al. classified their findings as 'spatial presence', 'involvement' and 'realness' whereas Lessiter et al. named theirs 'sense of physical space' (comparable to 'spatial presence'), 'engagement' (comparable to 'involvement') and 'ecological validity' (comparable to 'sense of physical space'). In addition, Lessiter et al. proposed a fourth category, called 'negative effects', that takes detrimental effects on presence into account.

Subject questionnaire measuring presence have difficulties accounting for variability in experimenter priming, subject-dependent confounds and inter-study comparisons. As a result, objective physiological measures such as heart rate and galvanic skin response have been tried with some success [6, 7]. These methods typically involve threatening stimuli such as an open, vertigo-inducing pit [8], or a verbal threat to damage a virtual limb [9]. Even though these measures can yield useful responses, the measurement infrastructure required can impair freedom of subject movement and adaptation effects can make it difficult to take more than one measurement per subject. Estimation of elapsed time has also been used to measure presence, albeit with inconclusive results [10, 11]. In a study by IJsselsteijn et al., subjects were required to subjectively estimate the speed of completion of a task, on a scale from one to six, and its actual duration in seconds. The results showed a positive correlation between presence and speed of task completion, but not estimated duration. In a study by Waterworth and Waterworth [11], subjects were asked to retrospectively estimate the time they spent in a media tent. Contrary to IJsselstein et al., Waterworth and Waterworth expected estimated times to be longer than the actual time spent in the media tent, but their results were inconclusive. Task enjoyment and mood also influences time perception: in a study of consumer behavior, positively valenced (major key) music caused subjects to overestimate elapsed time compared to negatively valenced music [12]. Other studies suggest that working memory influences time perception [13, 14], as does remembered knowledge of an event [15] and the tempo implicit in the language used to describe a past event [16]. However, the relationship between ownership presence and subject self-termination of a task after a fixed perceived time has not been tested. Self-termination time could offer a better chance of correlating highly with presence, because variations in individual recall capacities are removed from the measurement. Also, lower levels of presence will motivate the subject to abort the task sooner so that they can return to the real world where they (presumably) have more ownership of what is going on.

We hypothesize that a person's ability to estimate time is affected by his or her currently available attentional cognitive resources. This means that it should be more difficult in highly engaging virtual environments to accurately determine when a given time has elapsed, and that the passage of time should be underestimated (resulting in longer time spent in the virtual environment).

To test our hypotheses, we created four motor tasks with putatively different levels of immersion in a virtual environment. The levels of immersion were manipulated by varying the extent to which the tasks were goal-directed, and the degree of first-person control interactions available in the virtual environment, i.e. their ownership of objects in the environment.

We studied the relation between prospective time perception, task, ownership and presence, as assessed by subject performance in time estimation and in their responses to questionnaires. All procedures were approved by the Ethics Committee of the ETH Zurich.

2. Methods

2.1 Participants

58 subjects, mostly university students, were recruited for the study. The participants' ages ranged from 20 to 45 years (mean age 24.7; std. dev. 4.3,). The experiments were conducted in a single individual session of about 45 minutes for each subject. Subjects were paid USD 20 for their participation. Experiments were conducted around the middle of the day. All participating subjects gave written informed consent prior to the experiments and signed a receipt for having received remuneration. All written and verbal instructions were provided in the subject's choice of English or German.

2.2 Task and Material

Participants were seated comfortably at a desk on a height-adjustable chair in a quiet room, about 70 cm in front of a flat LCD TV monitor (Acer, 90 cm diagonal, 1366 x 768 pixels) (Figure 1). The monitor was connected to a PC (Dell Optiplex 745, 1 GB RAM). Custom-made data gloves with one 3D compass (Honeywell HMR3300) and three bend sensors (Abrams; thumb, index finger and middle finger) per hand were connected to the PC.

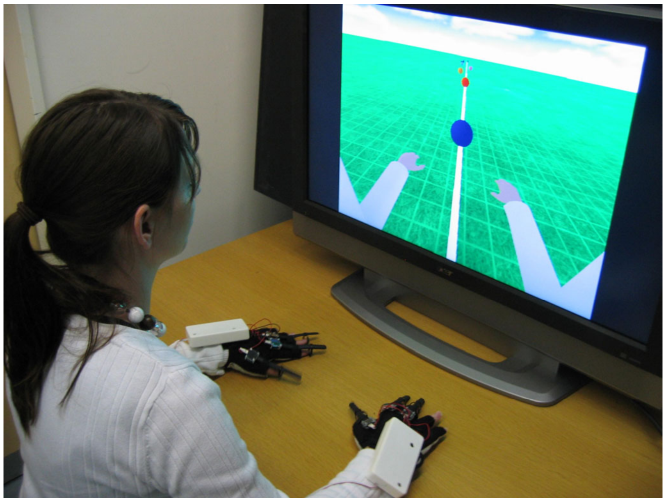

Fig. 1: The experimental setup with the data gloves and the screen with the presented interactive game

The monitor was used to display both the user instructions and the test scenarios. All test scenarios were based around a simple interactive game [17]. In the game, players have a first-person 3D view of two virtual arms and a large green field (Figure 1). The virtual arms are shown on the screen in front of the player in the same orientation and relative position as the player's real arms. The movements of the subjects' real arms are transferred to the virtual arms in real time. This close correspondence between the real and virtual arms in terms of position, relative orientation and movement was designed to induce the subject to treat the virtual arms as their own. When the game is played, different colored balls appear in the far distance and move along the field in a straight line towards the player. The trajectory of each ball is along a randomly chosen line directly towards the player, or some parallel offset away. The players' task is to move their arms to intercept the balls as they approach. Success or failure for each ball is indicated acoustically by different sound effects and visually if the ball is intercepted.

3. Experimental design

The experiment consisted of four tasks in a pseudo-randomized order determined by the experimenter. In each task the participant was instructed to start performing a certain activity at a 'go' signal and to tell the experimenter when they thought a minute had elapsed by saying 'stop'. Participants were instructed to not count the seconds in their head. They knew that the experiment consisted of four tasks, but they did not know the nature of the tasks or their order. All instructions were provided on-screen and orally elaborated to avoid confusion.

The four tasks were as follows:

- *Static screen task (SS):* Participants were instructed to freely move their hands on the table. The screen displayed a static picture of the arms in the default rest position.
- *Movement imitation task (MI):* Participants were instructed to imitate slow periodic inward-outward movements of the forearms in the horizontal plane executed by the virtual arms. The movements were displayed on-screen in a pre-recorded video sequence that ran until the participant stopped or a maximum of three minutes had elapsed.
- *Movement projection task (MP):* The participants were instructed to freely move their arms on the table. Their own movements were mapped directly onto the virtual arms on screen in real time.
- *Interactive game task (IG):* participants were instructed to play the interactive game, consisting of intercepting colored spheres rolling towards the player (as described above). Their arm movements were mapped onto the screen in the same way as for the movement projection task.

Two (MP and IG) of the four tasks included ownership components, and two tasks (SS and MP) allowed the participants to freely choose the movements they wish to perform.

To take possible sequencing effects into account, the tasks are given in four pseudo-random orders that were randomly assigned to participants, as summarized in Table 1. Every task appeared once in every position, and with a different task immediately following it.

Order	Task 1	Task 2	Task 3	Task 4
1	SS	MI	MP	IG
2	MI	IG	SS	MP
3	MP	SS	IG	MI
4	IG	MP	MI	SS

Table 1: Pseudo-random order of tasks. Each subject completed the tasks in one of these sequences. SS = static screen, MI = movement imitation, MP = movement projection, IG = interactive game.

4. Measurements and Analysis

Elapsed time was measured using a stop watch. After completing each task, the participants filled out a questionnaire. This break from the virtual environment forced the subjects to refocus every time a new task started. The questionnaire remained the same throughout the whole experiment, with the participants unable to see their answers from the previous tasks. Answers for questions two to five were given on a seven-point scale, with one indicating disagreement and seven indicating agreement. For the first question, a value of four indicated exactly 60 seconds, while three (-), two (--) and one (---) indicated too little time spent on the task, and five (+), six (++) and seven (+++) indicated too much time spent on the task on a relative scale. Participants were not informed of their time estimation performance between tasks or after the experiment. Table 2 shows the questionnaire given to the participants after each task.

#	Statement	Response scale (1-7)
1	How long do you think you actually performed the task? (Estimated Time)	--- -- - 60s + ++ +++
2	I found it easy to estimate the duration of time (1 min). (Ease of Duration Estimation)	disagree … agree
3	I enjoyed the motor task. (Enjoyment)	disagree … agree
4	It was easy for me to perform the required movements. (Ease of Movements)	disagree … agree
5	I felt immersed in the movements while performing them. (Presence)	disagree … agree

Table 2: Subject questionnaire given after each task

Statistical analysis was performed using SPSS with a mixed model of fixed and unfixed values. The fixed values comprised the ordering (i.e. the actual order in which the tasks were performed, see Table 1), the task itself (i.e. SS, MI MP, IG) and the sequence of the presented tasks (i.e. the role of the position in the sequence in which a given task is performed, independently of the task that preceded or followed it). In addition to the influence of the tasks themselves, this analysis allowed us to detect possible influences of the ordering of the presented tasks as well as general sequencing effects. For correct data analysis, we had to transform the actual times data logarithmically because the data distribution was not normal

5. Results

The results are summarized in Table 3 and Figure 2. The type of task had a highly significant effect on all investigated variables. Significant ordering and sequencing effects were found in the data. Order was significant for the items “ease of duration estimation” and “presence” and sequence was significant for “actual time” and “ease of movement”.

Measure	Task		Order		Sequence	
	F	p	F	p	F	P
Actual time	23.80	**0.000	1.39	0.255	6.10	**0.001
Estimated time	55.98	**0.000	1.46	0.231	0.87	0.353
Ease of duration estimation	22.51	**0.000	4.40	*0.040	0.10	0.320
Enjoyment	139.30	**0.000	3.17	0.080	1.63	0.203
Ease of movement	12.40	**0.001	0.12	0.915	7.49	**0.007
Presence	147.11	**0.000	6.24	*0.015	0.01	0.945

Table 3: F and p values of task, order and sequence for every investigated variable. * = 0.05 significance, ** = 0.01 significance.

On average, subjects spent significantly more than 60 seconds on all tasks. In general, more time was spent on the tasks which provided higher putative immersion; i.e. SS < MI < MP < IG. In the case of IG, the time was almost two minutes – a remarkably large error in time estimation. The overall time averages for SS, MI, MP and IG respectively are 71.9s ± 9.4s, 80.3s ± 10.6s, 78.6s ± 10.3s and 110.2s ± 14.5s (mean ± SE).

As shown in Figure 2, “actual time” increases with the putatively more presence evoking tasks. The same is true for “estimated times”. The generally low values for “ease of time estimation” indicate that the tasks were sufficiently absorbing to interfere with time estimation for all scenarios, from SS to IG. In addition, the data also suggests that time estimation became even harder when task was highly engaging. Reported enjoyment increases from SS to MI to MG to IG. For the “ease of movement” question, all values are high. Self-reported “presence” also increases for the putatively more interactive tasks.

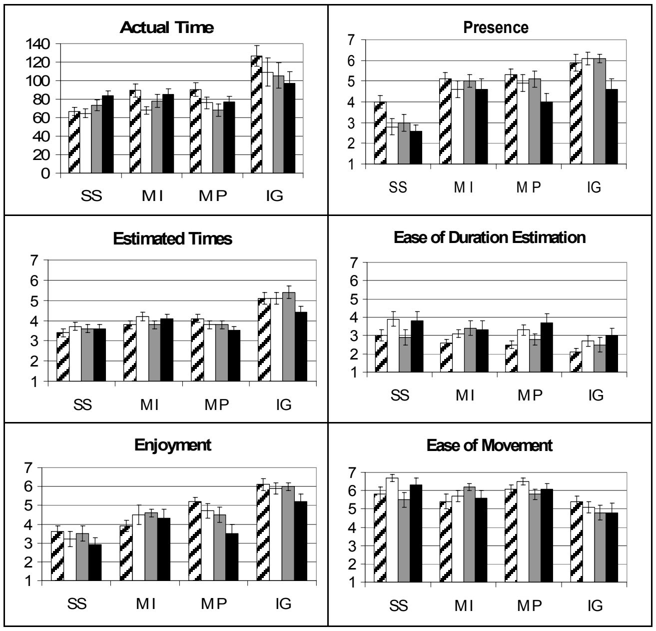

Figure 2: Mean results for each measurement, grouped by task and ordering. Error bars indicate ±1 standard error. Legend: striped bars = order 1, white bars = order 2, grey bars = order 3, black bars = order 4 (see Table 1).

6. Discussion

As we and IJsselsteijn [10] hypothesized, subjectively perceived time passes faster with the putatively more engaging tasks, leading to longer actual times spent on these tasks. This result suggests that prospective time estimation is a useful measure of presence in virtual reality. It is also an indicator of whether subjectively perceived times are under- or overestimated with respect to the actual amount of time passed. This finding contradicts the hypothesis of Waterworth and Waterworths [11] who postulated that the actual time spent should decrease with increasing immersion.

Significant ordering and sequencing effects were found in the data; i.e. subject performance in some measures depended on the order of task presentation, as well as the current time-point of the experiment independent of the task itself. Ordering effects can be explained by a conservative first choice for the questionnaire answers. Because participants knew that more tasks were to come, but they did not know their nature, they chose the first answer carefully, leaving enough room for higher or lower responses to later tasks. This caused answers for one order to generally be lower than for answers of other orders. For instance, the responses for “Presence” in Figure 2 for order number one are generally higher than for order number four. A baselining process removes this effect; i.e. subtract the value for the lowest rated task from all ratings for each order. However, even after baseling the influence of sequence remained for "actual time" and "ease of movement". We assume that these remaining sequencing effects are caused by a familiarization process for the "easy movement" question and a so-called time-order error for the actual time spent [13]. When two successive stimuli are to be rated with respect to time, the first one is often estimated shorter than the second one. According to Ornstein's [13] storage size hypothesis, this effect can be explained by a fading of information from the first sequence during the experience of the second sequence.

A sequencing effect was found for the actual time spent performing a task. Analysis of the graph of "actual time" in Fig. 2 suggests that more time is generally spent on later tasks than on earlier tasks. This means that if we had hypothetically presented the same task four times to the same person, the time spent per task would have increased from repetition to repetition. The effect was seen for our four different tasks, meaning that beside a task-dependency on the time spent, there was also a task-independent component which caused a trend for increasing time spent on a task. The cause of this sequencing effect is most likely the above mentioned time-order error.

The level of subjective immersion (in everyday usage of the term), i.e. presence (in our definition), depends both on the task and the assigned order. This finding can be explained by the fact that the tasks (from SS to MI to MP to IG) are not only gradually more engaging but can also be looked at as a natural training progression. For instance, the movements that subjects were told to imitate in the MI task were closely matched to the movements they had to perform in IG. In addition, the MP task probably gave participants a feeling of what it meant to project their real movements onto the screen. This progression assists in allowing the subjects to focus more intensely on the IG task when it is given. The smooth progression is not available in the other orderings, thus possibly causing reduced presence due to participants' unfamiliarity with the more complex tasks (IG, MP) when they are presented early in the ordering.

We expected the values for retrospectively estimated times spent in all four tasks to be constant, since it was a self-termination assignment. Surprisingly, however, participants did realize that they had spent qualitatively too much time on some tasks when asked to retrospectively estimate their actual times. This result suggests that retrospective time estimation is a poorer estimate of presence than self-termination of a task (prospective estimation), since subjects seemed to be able to guess that they had made a timing error after the fact. Other investigations [18, 19] have found that prospective and retrospective time estimation are inversely correlated; i.e. perceived time increases with increased cognitive load for retrospective time estimation, and decreases with increased cognitive load for prospective time estimation. Our data does not allow us to make a conclusive statement on this matter, as it contains effects from both the time-order error and the relation between prospective and retrospective time estimation. In addition, the repeated-measures protocol we used means that we cannot make meaningful correlations between the recorded variables.

It was not easy for participants to estimate the passing of time overall, and as expected, it was even more difficult for the more interactive tasks. Also, as expected, task enjoyment increased with the more engaging tasks.

Movements were rated as being easier for more precisely defined task (MI and IG), rather than for free movement tasks (SS and MP). The task may be easier for participants when they can follow instructions and do as they are told, without the burden of thinking about what kind of movement they want to carry out. Statistical analysis produced a significant sequencing effect which could possibly be explained by a familiarization process with the projection of the arm movements onto the screen.

Five of the 58 participants completely forgot about the time estimation task when performing the IG task and played until the game was finished (about 4 minutes). One of these five people had been randomly assigned to order 1, two to order 2, one to order 3 and one to order 4. This was an indication that the game task was immersive enough to cause about 8% of subjects to forget a more important task even in an unfamiliar lab setting, where subjects might be expected to be more tense and more focused than in their natural surroundings [20].

When we designed the experiments, the visual input was kept as similar as possible by using same display technology (monitor and speakers) and the same visual scenario in all cases, even though the tasks were obviously quite different. Of course, it is generally impossible to have exactly the same level of immersion for non-interactive and interactive scenarios, since the input technology used for interaction changes the level of immersion. We hypothesized that the time spent on the tasks would increase from SS to MI to MP to IG, with ownership having a big influence on perceived presence. However, the results for MI and MP were quite similar. This finding could possibly be explained by task-related differences. In the SS and MP tasks, participants were instructed to freely move their hands on the table in any way they wished. Some of the participants seemed unsure of how to move their hands, and some did not move them at all. Because participants could choose what kind of movement they wanted to perform in the SS and MP tasks, they had the freedom to become more or less involved in the task. For instance, participants who did not move their hands at all were most likely not involved at all. This means that the involvement was not uniform over all four tasks. Because ownership affects presence but not necessarily involvement, future experiments trying to dissociate the two should attempt to keep the ownership and task-related motivation effects constant.

Conclusions

Participants' surprisingly accurate retrospective time estimation suggests that a prospective, self-terminating task is more suitable for assessing subjectively experienced presence. However, a possible drawback of measuring presence using prospective time estimation could be that the attentional resources required to estimate the passing of time may take away resources from the subject's experience of presence. Our results suggest, however, that the required resources may be relatively independent of immersion levels, since the presence measures (timing and questionnaire responses) reflect the increases in immersion. Thus the resources required for prospective time estimation do not prevent the effective measure of presence.

Studies using prospective time estimation with repeated measures need to neutralize order-dependent task learning effects. In our protocol, it would have been beneficial to include a pre-experimental task-training phase to familiarize subjects with the tasks before testing.

To investigate presence independent of confounds caused by involvement, the ownership and task effects should be separated from each other. An experimental design that manipulates ownership without changing the task should include one case in which subjects are instructed to imitate movements performed by abstract elements, and a second case in which they imitate movements performed by virtual arms and hands. In general, though, the exclusion of involvement is difficult to achieve because inter-individual personality differences have to be taken into account. Other demographic factors such as gender, age, previous experience with virtual environments, and personal interests may also have an influence on individual involvement.

Acknowledgements

We thank Pawel Pyk, Edith Chevrier and Yves Hauser for designing the interactive game and the gloves used for these experiments, as well as Rahel Liesch for her support with the data analysis. This work was supported by the Swiss National Center of Competence in Neural Plasticity and Repair, and by the Gebert Ruef Foundation grant GRS-059-05.

References

[1] Slater, M. (2003) A Note on presence terminology. Presence-Connect

[2] Waterworth, E.L. and J.A. Waterworth, Focus, locus, and sensus: the three dimensions of virtual experience. Cyberpsychol Behav, 2001. 4(2): p. 203-213.

[3] Sas, C. and G.M.P. O'Hare, Presence equation: An investigation into cognitive factors underlying presence. Presence-teleop virt, 2003. 12(5): p. 523-537.

[4] Schubert, T., F. Friedmann, and H. Regenbrecht, The experience of presence: factor analytic insights. Presence-teleop virt, 2001. 10(No. 3): p. 266-281.

[5] Lessiter, J., et al., A cross-media presence questionnaire: The ITC-sense of presence inventory. Presence-teleop virt, 2001. 10(3): p. 282-297.

[6] Brogni, A., M. Slater, and A. Steed. More breaks less presence. in Presence 2003: The 6th annual international workshop on presence. 2003.

[7] Slater, M., A. Brogni, and A. Steed (2003) Physiological responses to breaks in presence: A pilot study. Presence 2003: The 6th annual international workshop on presence

[8] Meehan, M., et al., Physiological measures of presence in stressful virtual environments. ACM T Graphic, 2002. 21(3): p. 645-652.

[9] Ramachandran, V.S. and W. Hirstein, The perception of phantom limbs. Brain, 1998. 121: p. 1603-1630.

[10] IJsselsteijn, W., I. Bierhoff, and Y. Slangen-de Kort. Duration estimation and presence. in Presence Conference 2001. 2001. Philadelphia, PA, USA.

[11] Waterworth, J.A. and E.L. Waterworth, Being and time: Judged presence and duration as a function of media form. Presence-teleop virt, 2003. 12(5): p. 495-511.

[12] Kellaris, J.J. and R.J. Kent, The influence of music on consumers' temporal perceptions. J Consum Psychol, 1992. 1(4): p. 365-376.

[13] Ornstein, R.E., On the experience of time. Penguin Science of Behaviour. 1969, Baltimore, USA: Penguin Books.

[14] Fraisse, P., Perception and estimation of time. Annu Rev Psychol, 1984. 35: p. 1-36.

[15] Burt, C.D.B. and S. Kemp, Retrospective duration estimation of public events. Mem Cognit, 1991. 19(3): p. 252-262.

[16] Pedersen, A.C.I. and D.B. Wright, Do differences in event descriptions cause differences in duration estimates? Appl Cognitive Psych, 2002. 16: p. 769-783.

[17] Eng, K., et al. Cognitive virtual-reality based stroke rehabilitation. in World Congress on Medical Physics and Biomedical Engineering. 2006. Seoul, Korea: Springer.

[18] Hicks, R.E., G.W. Miller, and M. Kinsbourne, Prospective and retrospective judgments of time as a function of amount of information processed. Am J Psychol, 1976. 89(4): p. 719-30.

[19] Zakay, D. and R.A. Block, Prospective and retrospective duration judgments: an executive-control perspective. Acta Neurobiol Exp (Wars), 2004. 64(3): p. 319-28.

[20] Graf, P. and S. Grondin, Time perception in time-based prospective memory, in Timing the future: The case for a time-based prospective memory, J. Glicksohn and M.S. Myslobodsky, Editors. 2006, World Scientific Publishing. p. 1-24.