Stairway to Elders: Bridging Space, Time and Emotions in Their Social Environment for Wellbeing

Giuseppe Boccignone[1] [a], Claudio de'Sperati[2] [b], Marco Granato[1] [c], Giuliano Grossi[1] [d], Raffaella Lanzarotti[1] [e], Nicoletta Noceti[3] [f] and Francesca Odone[3] [g]

[1]*Department of Computer Science, Università degli Studi di Milano, Milan, Italy*

[2]*Department of Psychology, Università Vita-Salute San Raffaele, Milan, Italy*

[3]*MaLGa Center - DIBRIS - Università di Genova, Genova, Italy*

Keywords: Emotion Recognition, Social Attitude, Video Speed Tuning, Emotion Regulation, Emotion-aware Ambient Intelligence (AmE), Multi-modal Data, Gerontechnology.

Abstract: The physical and mental health in elderly population is an emergent issue which in recent years has become an urgent socio-economic phenomenon. Computer scientists, together with physicians and caregivers have devoted a great research effort to conceive and devise assistive technologies, aiming at safeguarding elder health, while a marginal consideration has been devoted to their emotional domain. In this manuscript we outline the research plan and the objectives of a current project called *Stairway to elders: bridging space, time and emotions in their social environment for wellbeing"*. Through a set of sensors, which include cameras and physiological sensors, we aim at developing computational methods for understanding the affective state and socialization attitude of older people in ecological conditions. A valuable by-product of the project will be the collection of a multi-modal dataset to be used for model design, and that will be made available to the research community. The outcomes of the project should support the design of an environment which automatically (or semi-automatically) adapts its conditions to the affective state of older people, with a consequent improvement of their life quality.

1 INTRODUCTION

Nowadays, especially in developed countries, life expectancy keeps growing, augmenting the proportion of older people over the population. According to the WHO, the worlds elderly population (defined as people aged 60 and older) has increased drastically in the past decades and will reach about 2 billion in 2050. In Europe, the percentage of the EU27 population above 65 years of age is foreseen to rise to 30% by 2060 (Giannakouris et al., 2008). This evidence has opened a social debate about how to face this socio-economic phenomenon.

Indeed, the human being is a complex organism, whose wellbeing may be described following several dimensions, encompassing the physical, psychological, economic, and social domains. The process of aging typically reduces the individuals potential in one or more of these domains, leading to a condition of vulnerability and clinical instability, defined as frailty in the specialized literature (Fried et al., 2001). In the last decades, a great research effort has been devoted to the design of assistive technologies that have positive impacts on different dimensions of health and quality of life for aging population (Yared and Abdulrazak, 2016). Also, the concept of smart environments design has opened the possibility of monitoring patients at home (Scanaill et al., 2006). To date, a vast majority of methods proposed in this field addresses the problem of monitoring health status of people mainly considering physical attributes while the emotional and social domains have received only marginal consideration.

Emotional/affective wellbeing is deeply associ-

[a] https://orcid.org/0000-0002-5572-0924
[b] https://orcid.org/0000-0002-7322-2240
[c] https://orcid.org/0000-0002-7322-4350
[d] https://orcid.org/0000-0001-9274-4047
[e] https://orcid.org/0000-0002-8534-4413
[f] https://orcid.org/0000-0002-6482-4768
[g] https://orcid.org/0000-0002-3463-2263

Boccignone, G., de'Sperati, C., Granato, M., Grossi, G., Lanzarotti, R., Noceti, N. and Odone, F.
Stairway to Elders: Bridging Space, Time and Emotions in Their Social Environment for Wellbeing.
DOI: 10.5220/0009106605480554
In *Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2020)*, pages 548-554
ISBN: 978-989-758-397-1; ISSN: 2184-4313

ated with the sense of fulfilment, including satisfaction, optimism, having a purpose in life as well as being able to make the most of your abilities to cope with the normal challenges of life. As suggested in (Hawkins, 2005), the mental wellbeing of the aging population is, together with health, a key factor for aging well since, besides irrefutable physical needs, the wellbeing of a person cannot leave out emotional and social aspects (Anttonen and Surakka, 2007). Unfortunately, psychological stability is often undermined over the years, when older people have to face life challenges that weaken their independence and self-confidence.

These considerations are at the core of our current project called *Stairway to elders: bridging space, time and emotions in their social environment for wellbeing"*, aiming at devising a framework to enhance the condition of wellbeing for the aging population, focusing on affective, cognitive, and social factors.

In this note the project is introduced by outlining the main research ambition, namely the development of computational methods for understanding the affective state and the socialization attitude of older people. This will be grounded on multimodal observations acquired over medium/long temporal frames. In order to reduce the Hawthorne effect, that is the participant's behaviour alteration because of the awareness of participating in an investigation (Jones, 1992), the experiments will be carried out in unconstrained ecological environments, targeting the elderly common activities. In cascade, we will explore specific interventions on the environment able to affect positively on the individual mood.

The paper is organized as follows: in Sec. 2, we present a brief overview of the most recent Emotion-aware Ambient intelligence; in Sec. 3, we provide a detailed outline of our research objectives. Then, in Sec. 4, we illustrate the expected results, and in Sec. 5 we draw some conclusions and provide final considerations.

2 RELATED WORKS

Most of the investigations conducted in affective computing have been performed in reductive contexts, which may limit their applicability and may generate biased outcomes (Jallais and Gilet, 2010; Zhang et al., 2014). On the contrary, our project is grounded on the idea of implementing it in ecological environments, enriching the so called Ambient Intelligence (AmI) with actuators, so as to make it responsive to human needs and emotions. This concept is often referred to as Emotion-aware Ambient intelligence (AmE) (Zhou et al., 2007).

Some proofs of concept have already been presented in the literature, even though they do not propose specific computational models for elderly wellbeing. In (Fernández-Caballero et al., 2016), for example, the authors propose a generic, open and adaptable AmE architecture for emotion detection and regulation. This architecture should capture and integrate physiological signals, face expression and body movement and estimate the emotional state of the monitored person. A pleasant state of mood should then be induced by music and colour/light actuation. In (Mano et al., 2016) a three layer AmE architecture is introduced relying upon the following: simple and dedicated devices that act as sensors collecting the information about the patients health/emotional status; a decision maker with more powerful computing resources; an actuator, that should be a simple alert to caregivers. In (Rodrigues and Pereira, 2018), a unified model promoting the wellbeing of the elderly living at home is proposed. It takes into account three aspects concerning the wellbeing: health, activity, and emotions. As for emotions, they adopt a smiling count to determine the happiness of people. Such measure would then be integrated with personalized health and activity coefficients so that the AmE could activate proper actuators (e.g. to propose an activity, turn on a display, or alert the caregivers).

Overall, the most crucial lack we observe in existing solutions is that the detection of affective state and social attitude - namely, the core of AmE - is often obtained referring to general purpose classifiers taken on-the-shelf from generic affective computing tools (Grossi et al., 2019). Here and in our project we point out the importance of learning techniques relying on realistic data acquired in ecological conditions and involving older people, allowing to tailor a solution for this peculiar population.

3 RESEARCH FLOW

Goal of this project is to design a framework endowing it with the automatic assessment of affective state and social attitude in the elderly population; the model should include a feedback to allow possible corrective interventions to spatio-temporal environment, in case low-mood or negative attitude are observed (see Fig. 1).

Specifically, the project concerns three main objectives: (1) Identification of key contexts and the consequent acquisition of ecological data through multimodal and non-invasive sensors (e.g., cameras, and wristbands), covering an appropriate space and time, and promoting comfort during the acquisition. (2) Definition and implementation of models for the quantitative evaluation of the affective state and the social attitude of an individual in an ideal continuum of the emotional space, with particular attention to subtle changes (as an episode of anger or frustration due to a temporary failure in a performed activity) or, conversely, slower variations that can be only appreciated through monitoring over longer time periods (e.g. an increasing sense of unease due to an unpleasant social context). (3) Provision of feedback, with mild interventions on the space and time dimensions (to improve comfort and stimulate positive feelings), and an eventual customization of the intervention, with respect to a specific person or a particular mood-state.

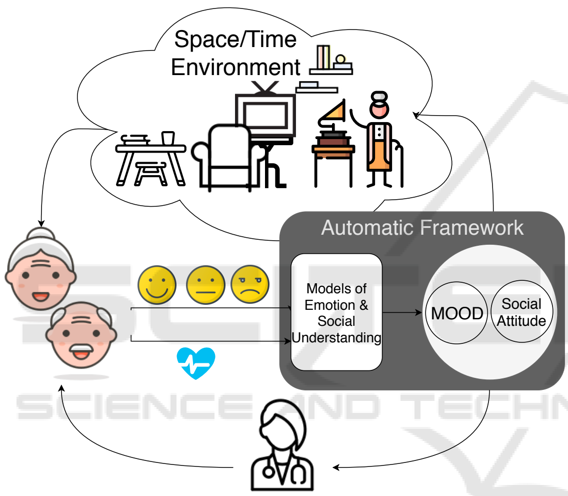

Figure 1: Workflow of the proposed system: the environment influences both emotions and social interactions of elderly people; social signals are acquired and an automatic framework provides the social attitude level and the emotional state so as to guide contingent interventions.

In the following we comment on each proposed objective.

3.1 Key Contexts

The identification of suitable monitoring scenarios is a key element to spot possible frailties. As detailed in the following, in our project we will take into account normal activities of daily living in a homely environment, one-to-one laboratory sessions aimed at stimulating motor or cognitive skills, and social activities that require either interaction among elders or with younger people (e.g. family members, operators, visitors). Such a variety of contexts is possible thank to the collaboration with different stakeholders allowing us to have available multiform environments, several live styles (in community or in private flats), and different degrees of frailty. In particular, we will involve elders attending a university of the third age (minimal frailty), elders living in a rest home (with multiform level of frailties), and others living in flats annex to a hospital (various level of frailties).

3.1.1 Watching Television

For aged people living alone, watching TV often represents one of the main daily activities, thus deserving a specific study to make this time fruitful in giving a positive feeling to elders, and to investigate the factors to favor this. For example, we will compare the effect of conducting such activity alone or in a group and the effect of tuning the video speed to users perceptual and cognitive abilities. Assessing the core affect of the TV watcher (via physiological data, face expression, and questionnaires) will allow us to derive his/her mood, and consequently to intervene on the fruition modality to induce a sense of fulfilment.

3.1.2 Free Time

This context concerns all the unconstrained scenarios, where elders conduct activities moving in the space freely (e.g. reading, playing board games, having a conversation or taking a rest, cooking). This context is suitable to reveal the way elders spontaneously relate with the environment around them, their psychological stability and, if conducted in a social context, their socialization attitude. In this task, we will investigate the physiological data and body movement, more reliable than other cues given the unconstrained scenario. Specific attention will be posed on designing motion representations able to embed emotional features, bridging well known formulations describing biological motion and its regularities (Noceti et al., 2017) with the popular valence-arousal model (Russell, 2003).

As for the socialization attitude, it will be investigated through a videobased approach analysing groups formation and the mutual relation between people in the group, exploiting pose (Cao et al., 2018) and gaze direction (Recasens et al., 2015; Dias et al., 2019) estimation. The derived information will guide automatic changes in the environmental conditions (light and music), in order to facilitate emotion regulation.

3.1.3 Laboratory

This group of experiments will address specifically community life, where operators propose and conduct individual or group activities to stimulate either cognitive/affective, motor or social aspects. Information gathered within-lab, concerning the positive or negative impact on subjects, provides an essential feedback to the operators for understanding whether the stimulation is effective or not, and in which measure. In such a controlled setting, we will gather physiological signals, body movements, and, if possible, face expressions. Also, in this task, we will investigate the mutual pose and gaze direction, finalized to explore the degree of involvement of older people in the activity, and also their mutual collaboration.

3.1.4 Video Call

This task will concern the video call between elders and relatives or friends. A positive communication will concern both emotional and cognitive aspects. On the one hand, monitoring the emotion aroused by accomplishing this task will allow to assess the psychological stability of elders. On the other hand, we are interested in exploring the possibility of tuning the video flow according to the preferences of both speakers and listeners (typically slower for elders), for example by introducing a buffer in the communication or by slowing down pitchcompensated video communication.

3.2 Emotion and Social Attitude Quantification

This module aims at determining in a continuum the emotional state and social attitude of elders, by using multimodal information, i.e., facial expression, physiological signals, and motion quality. This is useful both when an individual is alone, to allow automatic environment changes or to give alert to relatives, and when he/she lives in a community, where the professional operators might miss subtle discomfort or poorly expressed emotions or needs. This objective will be accomplished facing three tasks, namely the emotion recognition, the social attitude estimation, and the analysis concerning video tuning.

3.2.1 Core Affect Learning from Multimodal Signalling

In order to exploit, when possible, the strength of multimodal social signals, a probabilistic framework (Boccignone et al., 2018) will be exploited, suitable for dealing with the variety of input signals and related uncertainties. Such a model accounts for multimodal signal dynamics in terms of trajectories unfolding in a continuous, central affect state-space (cfr., (Anderson and Adolphs, 2014) for an in-depth discussion) akin to the well known valence-arousal core affect space proposed by Russell (Russell, 2003) (see Fig. 2).

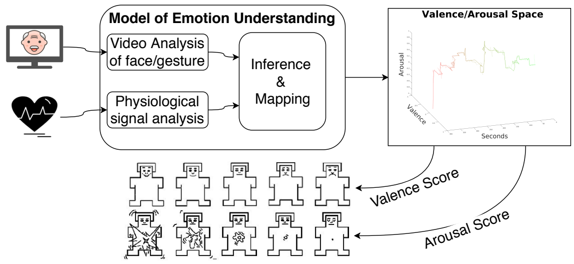

Figure 2: Workflow of the emotion sensing and understanding: social signals are captured and mapped into the valence/arousal space by a Bayesian deep model.

We will apply the model to a combination of physiological and facial cues, and consider the applicability to motion features and fullbody expressions.

The model is to be learnt in a supervised setting, from an appropriate amount of data, labelled (Boccignone et al., 2017) by domain experts according to the perceived emotion, either producing a personalized model for domestic use, or a general one for communities.

3.2.2 Social Attitude

Social attitude will be derived analysing gaze direction (Dias et al., 2019; Cuculo et al., 2018), body pose (Cao et al., 2018), and movements of people forming a group. The literature in expressive motion analysis is largely based on motion capture systems, with fewer results on video analysis. Video signals are instead easier to adopt (cheaper sensors, lower setup requirements) and they provide rich sources of heterogeneous information. For this reason, they will be considered as a main source of data in our project; this is feasible thanks to the availability of recently proposed pose estimation methods, which allow us to obtain information on the 2D or 3D body posture from video. We will first evaluate the applicability of the existing algorithms to the specific application domain, possibly improving, if necessary, the pose estimation precision. We will also incorporate a temporal analysis of the pose, based on dynamic filters. Then, we will exploit these measurements to obtain an estimate of the quantity and quality of social interactions among subjects, as a further cue of the overall affective state. As already stated, the motion features and full-body expressions will also provide a first coarse estimate of the emotional state, (extending (Piana et al., 2016)) that could be used as an additional input for the multimodal model described in Sec. 3.2.1.

3.2.3 Video Tuning

This task aims at defining the optimal tuning to video speed and rhythms, and at finding possible relations with individual sensorimotor rhythms, which in the elderly may be rather slowed (Salthouse, 1996). We will investigate video speed preferences at the perceptual and emotional levels with psychophysical and psychophysiological methods (Rossi et al., 2018; Mackersie and Kearney, 2017; deSperati and Baud-Bovy, 2017; Zuliani et al., 2019; de'Sperati, 2020)

and facial emotion recognition. We will keep into account individual abilities in the cognitive (time estimation through duration reproduction tasks (Grondin, 2010)) behavioural (through the analysis of speed-accuracy tradeoff, (Heitz, 2014)) and motor (through kinematic analysis of simple movements such as grasping (Bruno et al., 2016)) domains, as well as contextual factors such as time of the day, arousal, mood, etc. The expected outcome is the comprehension of how to optimally regulate video speed (increased speed, decreased speed, adjustable speed control, no change), which in perspective may lead to a change in TV and video standards to meet not only the elder population.

3.3 Intervention Definition

Given the assessment of emotional state and social attitude, several interventions could be considered, all finalized to enhance the elders life condition and gain emotional stability (Jallais and Gilet, 2010; Zhang et al., 2014). The interventions will be either automatic or guided by operators.

The automatic interventions on the domestic space will include changes in sensory stimulation such as light modulation (Cuculo et al., 2015) or music stimulation (Anttonen and Surakka, 2007), the former to adjust the environment in order to match the relaxation/arousing state or to favour interaction, and the latter to modulate or evoke emotions. Furthermore, the regulation of the reproduction speed of certain TV programs, and of the video flow during video call will induce a sense of fulfilment.

Finally, in close collaboration with the stakeholders and the caregivers, we will identify interventions guided by operators, aiming at improving the emotional state and the sense of inclusion.

4 EXPECTED OUTCOMES AND SOCIAL IMPACT

Our proposal answers to the emerging question on how to improve the quality of life of elders, (Martini et al., 2018), considering in particular their emotional and social spheres. Overall, the research will promote the design of protocols for the longterm, automatic, and quantitative assessment of the emotional well-being level of aging population, to complement the traditional evaluations performed by the physicians. Also, it will explore the possibility of producing adaptive feedback or interventions to the surrounding environments, to control and improve such emotional state. We identify in particular the following general aspects the project can bring impact on:

- the research will provide a stream of analysis to evaluate the overall emotional status of an aged person.
- It will introduce strategies to "represent" the personal perception of space and time of an individual, with consequent understanding of his/her sense of loneliness and social attitude.
- The framework will be designed to be reliable in social care facilities as well as private homes, providing an adaptive solution to generic settings and a variety of social contexts.
- An affective dataset focused on older people will be released, including physiological data, facial expression descriptors (e.g. Action Units (Ekman and Friesen, 1971)), and emotions annotations on the valence-arousal 2-dimensional space. This dataset will be freely available, as benchmark, to the scientific community.

5 CONCLUSION

In this paper, we highlighted the emergent issue of accounting for the psychological wellbeing of elder people, outlining a proposal to address the hitherto neglected problem of estimating elderly wellbeing in ecological or close-to-ecological contexts. Indeed, besides the physical wellbeing, current health facilities do not actually focus on mental and social aspects, that should deserve equal attention.

Under such circumstances, the rationale behind our proposal is to conceive a well founded framework to automatically assess the affective state in aging population in order to promote suitable intervention. Namely, the research activity will focus on the design and validation of methods for continuous detection and analysis of the emotional wellbeing and social interaction level of aging people, from multimodal sources of information in an ecological setting. Meanwhile, monitoring activity will be paired with strategic interventions deployed from the environment in order to induce positive emotions, in case of lowmood detection, and to improve the quality of social interactions.

ACKNOWLEDGEMENTS

This work has been supported by Fondazione Cariplo, through project Stairway to elders: bridging space, time and emotions in their social environment for wellbeing, grant no. 2018-0858.

REFERENCES

Anderson, D. J. and Adolphs, R. (2014). A framework for studying emotions across species. *Cell*, 157(1):187–200.

Anttonen, J. and Surakka, V. (2007). Music, heart rate, and emotions in the context of stimulating technologies. In *International Conference on Affective Computing and Intelligent Interaction*, pages 290–301. Springer.

Boccignone, G., Conte, D., Cuculo, V., DAmelio, A., Grossi, G., and Lanzarotti, R. (2018). Deep construction of an affective latent space via multimodal enactment. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4):865–880.

Boccignone, G., Conte, D., Cuculo, V., and Lanzarotti, R. (2017). Amhuse: a multimodal dataset for humour sensing. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pages 438–445. ACM.

Bruno, N., Uccelli, S., Viviani, E., and de'Sperati, C. (2016). Both vision-for-perception and vision-for-action follow weber's law at small object sizes, but violate it at larger sizes. *Neuropsychologia*, 91:327–334.

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., and Sheikh, Y. (2018). Openpose: realtime multi-person 2d pose estimation using part affinity fields. *arXiv preprint arXiv:1812.08008*.

Cuculo, V., DAmelio, A., Lanzarotti, R., and Boccignone, G. (2018). Personality gaze patterns unveiled via automatic relevance determination. In *Federation of International Conferences on Software Technologies: Applications and Foundations*, pages 171–184. Springer.

Cuculo, V., Lanzarotti, R., and Boccignone, G. (2015). The color of smiling: computational synaesthesia of facial expressions. In *International Conference on Image Analysis and Processing*, pages 203–214. Springer.

de'Sperati, C. (2020). Cortical resonance to visible and invisible visual rhythms. *Brain Sciences*, Submitted.

deSperati, C. and Baud-Bovy, G. (2017). Low perceptual sensitivity to altered video speed in viewing a soccer match. *Scientific reports*, 7(1):15379.

Dias, P. A., Malafronte, D., Medeiros, H., and Odone, F. (2019). Gaze estimation for assisted living environments. In *WACV*.

Ekman, P. and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124.

Fernández-Caballero, A., Martínez-Rodrigo, A., Pastor, J. M., Castillo, J. C., Lozano-Monasor, E., López, M. T., Zangróniz, R., Latorre, J. M., and Fernández-Sotos, A. (2016). Smart environment architecture for emotion detection and regulation. *J. Biomed. Inform.*, 64:55–73.

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., et al. (2001). Frailty in older adults: evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3):M146–M157.

Giannakouris, K. et al. (2008). Ageing characterises the demographic perspectives of the european societies. *Statistics in focus*, 72:2008.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, 72(3):561–582.

Grossi, G., Lanzarotti, R., Napoletano, P., Noceti, N., and Odone, F. (2019). Positive technology for elderly well-being: A review. *Pattern Recognition Letters*.

Hawkins, B. A. (2005). Aging well: toward a way of life for all people. *Preventing chronic disease*, 2(3).

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in neuroscience*, 8:150.

Jallais, C. and Gilet, A.-L. (2010). Inducing changes in arousal and valence: Comparison of two mood induction procedures. *Behavior research methods*, 42(1):318–325.

Jones, S. R. (1992). Was there a hawthorne effect? *American Journal of sociology*, 98(3):451–468.

Mackersie, C. L. and Kearney, L. (2017). Autonomic nervous system responses to hearing-related demand and evaluative threat. *American journal of audiology*, 26(3S):373–377.

Mano, L. Y., Faiçal, B. S., Nakamura, L. H., Gomes, P. H., Libralon, G. L., Meneguete, R. I., Filho, G. P., Giancristofaro, G. T., Pessin, G., Krishnamachari, B., and Ueyama, J. (2016). Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition. *Comput. Commun.*, 89:178–190.

Martini, C., Barla, A., Odone, F., Verri, A., Cella, A., Rollandi, G. A., and Pilotto, A. (2018). Data-driven continuous assessment of frailty in older people. *Frontiers in Digital Humanities*, 5:6.

Noceti, N., Odone, F., Sciutti, A., and Sandini, G. (2017). Exploring biological motion regularities of human actions: A new perspective on video analysis. *ACM Transactions on Applied Perception (TAP)*, 14(3):21.

Piana, S., Staglianò, A., Odone, F., and Camurri, A. (2016). Adaptive body gesture representation for automatic emotion recognition. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 6(1):6.

Recasens, A., Khosla, A., Vondrick, C., and Torralba, A. (2015). Where are they looking? In *Advances in Neural Information Processing Systems*, pages 199–207.

Rodrigues, N. and Pereira, A. (2018). A user-centred well-being home for the elderly. *Applied Sciences*, 8(6):850.

Rossi, F., Montanaro, E., and deSperati, C. (2018). Speed biases with real-life video clips. *Frontiers in Integrative Neuroscience*, 12:11.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, 110(1):145.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological review*, 103(3):403.

Scanaill, C. N., Carew, S., Barralon, P., Noury, N., Lyons, D., and Lyons, G. M. (2006). A review of approaches to mobility telemonitoring of the elderly in their living environment. *Annals of biomedical engineering*, 34(4):547–563.

Yared, R. and Abdulrazak, B. (2016). Ambient technology to assist elderly people in indoor risks. *Computers*, 5(4):22.

Zhang, X., Yu, H. W., and Barrett, L. F. (2014). How does this make you feel? a comparison of four affect induction procedures. *Frontiers in psychology*, 5:689.

Zhou, J., Yu, C., Riekki, J., and Kärkkäinen, E. (2007). Ame framework: a model for emotion-aware ambient intelligence. In *Proc. of the Int. Conf. on Affective Computing and Intelligent Interaction*, page 45.

Zuliani, E., Caputi, M., Scaini, S., and de'Sperati, C. (2019). Videos look faster as children grow up: sense of speed and impulsivity throughout primary school. *Journal of experimental child psychology*, 179:190–211.