PROCEEDINGS OF THE 16TH SWECOG CONFERENCE

Virtual from Skövde 2021

10 - 12 November

Editors

Erik Billing
Andreas Kalckert

1

SweCog
SWEDISH COGNITIVE SCIENCE SOCIETY

Editors:
Erik Billing, University of Skövde, erik.billing@his.se
Andreas Kalckert, University of Skövde, andreas.kalckert@his.se

www.swecog.se

The SweCog 2021 conference was financially supported by:
The School of Informatics and The School of Bioscience at the University of Skövde.

Skövde University Studies in Informatics 2021:2
ISBN 978-91-983667-8-5
ISSN 1653-2325

PUBLISHED BY THE UNIVERSITY OF SKÖVDE

Preface

We welcome you to the 16'th SweCog conference! After the 2020 meeting had to be cancelled, due to the unusual circumstances of facing a worldwide pandemic, we look forward to finally meet again, although the pandemic makes us meet virtually and not in person.

Fittingly, an emerging theme of this year's meeting is virtual reality. A technology which creates new ways of interacting with each other and with the world. It is not only a subject of active research, but increasingly also a medium for new creative experiments or applications, as evidenced by one of our keynote speakers this year. VR has become now a more widely available tool in different areas of research, and probably has made its full and final impact not yet.

SweCog 2021 also features a nod to the *word usability day*. As technology becomes increasingly present in our daily lives, not the least emphasized through the pandemic, we believe that cognitive science has an important role as a field of research informing the design of usable digital artifacts. As the University of Skövde stands as one example of the close relation between cognitive science and user experience design, we take the opportunity to celebrate the topic of *Cognitoon and UX*.

This meeting has been organized jointly by the Interaction lab and the Cognitive Neuroscience lab of the University of Skövde. We are glad to see this interaction happening between the two labs and the two fields. We hope this is not perceived as an "invasion" of the brain scientists documenting the failure of cognitive science as a field (see Nunez et al., 2019), but rather a collaborative move of finding synergies in our research. In this spirit, we hope our meetings continue to bring people together from different parts of Sweden, from different departments, and maybe also from more different disciplines, to discuss our latest research. And despite our enthusiasm for virtual reality, we sincerely hope the next meeting will allow us to meet again in person.

Erik Billing and Andreas Kalckert

The reviewers were:

Beatrice Alenljung, Erik Billing, Linus Holm, Arne Jönsson, Andreas Kalckert, Erik Lagerstedt, Maurice Lamb, Kajsa Nalin, Jonas Olofsson, Anders Persson, Jana Rambusch, and Katja Valli.

Conference Programme

Wednesday November 10th

13:00 — 13:05	*Welcome*
13:05 — 14:05	Invited speaker - **Mel Slater** (University of Barcelona, see p. 7)
14:20 — 14:40	Raphael Aybar (p. 15)
14:40 — 15:00	Gordana Dodig-Crnkovic (p. 22)
15:00 — 15:20	Philip Millroth (p. 6)
15:30 — 15:50	Christian Balkenius (p. 19)
15:50 — 16:10	Joel Parthemore (p. 35)
16:10 — 16:30	**Poster talks** (see p. 9)
16:30 — 17:30	**Poster session**

Thursday November 11th

13:00 — 13:05	*Welcome*
13:05 — 13:25	**Jana Rambusch** - Introduction to Cognition and UX
13:25 — 13:45	Kata Szita (p. 41)
13:45 — 14:05	Fateme Teimouri (p. 44)
14:20 — 15:20	Invited speaker - **Virpi Roto** (Aalto University, see p. 6)
15:30 — 15:50	Axel G. Ekström (p. 26)
15:50 — 16:20	**Poster talks** (see p. 9)
16:20 — 17:30	**Poster session**

Friday November 12th

13:00 — 13:05	*Welcome*
13:05 — 14:05	Invited speaker - **Agneta Gulz** (Lund University & Linköping University, see p. 6)
14:20 — 14:40	Paul Hemeren (p. 29)
14:40 — 15:00	Claes Strannegård (p. 38)
15:00 — 15:20	Maurice Lamb & Rebecca Rouse (p. 32)
15:20 — 15:30	*Final remarks*

Verbal presentations

Below you'll find abstracts for the talk by Philip Millroth and for the invited key notes by Agneta Gulz, Virpi Roto, and Mel Slater. In addition, the following short papers are presented verbally:

Agneta Gulz — Introducing learning science to teacher students at Swedish universities

In the presentation I will go through the topics from my two introductory lectures on "learning science" for teacher students. The lectures, targeting students at the end of their first year, have so far only been given one time; thus there is much leeway to modify and change them. I look forward to ideas and comments from the audience, not least those who have experiences from teacher education.

Philip Millroth — Toward a Richer Understanding of Human Cognition: Unleashing the Full Potential of the Concurrent Information-Processing Paradigm

One of the most influential working hypotheses in psychology, to this day, is that human information processing in higher-order cognition (e.g., judgment and decision-making) is constrained by having to process objects serially, one at a time. However, a rather large body of research, accumulated over the past 50 years, has demonstrated that serial-processing models provide a poor descriptive account of human information processing. An alternate to the serial-processing view is that people can process information concurrently; many cognitive processes can advance independently of each other even if the system involves only a single central information-processing unit (i.e., central executive). Perhaps this general idea can be advanced beyond its present standing and is to provide new powerful tools in the pursuit of understanding human behavior. To this end, the present study provides a review of (i) the conceptual differences between different types of processing (serial, concurrent, parallel), (ii) recent advancements in the field of computer science, and (iii) existing research on human information-processing, which is in line with the advancements in computer science. Finally, the study provides a discussion of outstanding research questions gleaned from these reviews—questions that could stimulate entirely new research programs.

Virpi Roto — Is this User Experience Design?

Human-Computer Interaction, the home field of user experience research, has evolved over 40 years from cognitive aspects to usability, to user experience, and lately to well-being. Today, good user experience is a top priority in companies developing digital services for consumers, but the interpretation of what user experience design means remains limited. This is the case especially in work contexts where good user experience is often interpreted in simplistic way as ease and efficiency of use. In this talk, I will dig into the concept of user experience, the status of user experience research in work context, and report examples of our experience design research at work. Finally, I look at the future and how work automation changes Human-Computer Interaction research. With intelligent systems, we are facing a major challenge of redefining the design goals for interaction on many levels, from cognitive to affective experience.

Mel Slater — Body Representation in Virtual Reality

How the brain represents the body is a fundamental question in cognitive neuroscience. Experimental studies are difficult because the body is always there (William James). In recent years immersive virtual reality techniques have been introduced that deliver apparent changes to the body extending earlier techniques such as the rubber hand illusion, or substituting the whole body by a virtual one visually collocated with the real body, and seen from a normal first person perspective. This talk will introduce these techniques, and concentrate on how changing the body can change the mind and behaviour, especially in the context of combatting aggression based on gender or race.

Posters

Peter Bang, Maria Strömberg & Kajsa Igelström — Associations between self-reported sensory sensitivity and social, communicative, and rigid autistic traits

Background: Autism spectrum conditions have long been associated with abnormal sensory processing, but it is not known whether problems within specific sensory modalities are differentially associated with specific autistic traits. In this study, we examined the relationship between sensory sensitivity and autistic traits. Sensory sensitivity was measured using the modality subscales of the Glasgow Sensory Questionnaire, and autistic traits in social, communicative, and rigid domains were measured using the Broad Autism Phenotypic Questionnaire. Methods: The questionnaires were administered using Qualtrics, and recruitment of participants was handled using Prolific. The total cohort was N=252 consisting of both autistic and neurotypical adults. Results: We found that autistic traits in all three domains were predicted by sensitivity in the auditory modality. The communicative and rigid domains were furthermore predicted by gustatory sensitivity, while sensitivity in the tactile modality was specific for the social domain.

Nikolaos Chrysanthidis — Episodic memory semantization in a spiking cortical memory model

Episodic memory (EMs) is susceptible to loss of information [Tulving, 1972], which can be partly accounted for by semantization. Extensions to the classical Remember/Know behavioral paradigm attribute episodic forgetting to repeated exposures of memory items in different contexts leading to semantization [Opitz, 2010]. Since this prominent phenomenon has not been offered any solid mechanistic explanation, we propose and evaluate a Bayesian-Hebbian hypothesis about synaptic and network factors contributing to the effect of EM semantization. In particular, we build a model of two cortical spiking neural networks associatively coupled using a Bayesian-Hebbian learning rule (BCPNN) [Tully et al., 2014, Fiebig et al., 2020], and show how it captures key phenomenological aspects of the semantization. We simulate an EM task designed to follow a seminal experimental study [Opitz, 2010], and qualitatively compare the modelling results with the corresponding behavioral data. We demonstrate that encoding items across multiple contexts leads to item-context decoupling akin to semantization. The emerging loss of episodicity progresses with further exposures of a stimulus in different contexts, resulting in weaker item-context memory binding. This gradual trace modification relies on the nature of Bayesian learning dependent on pre- and post-synaptic activity. Our model bridges neural mechanisms with behavioral outcomes, and reproduces important EM phenomena abiding by various biological constraints.

Acknowledgments

Vetenskapsrådet 2018-05360; SNIC resources at PDC Centre for High Performance Computing, KTH Royal Institute of Technology; Swedish e-Science Research Centre (SeRC); KTH Digital Futures.

Adam Enmalm — The role of visual perspective on self-touch perception: An exploratory study of somatosensory attenuation

Whenever you touch yourself, the ensuing touch feels distinctly different from when someone else touches you. Self-generated touch is perceived as less intense, a phenomenon called somatosensory attenuation. Touch to the self, regardless of origin, is perceived through a first-person perspective, whereas touch to another is perceived through a third-person perspective. The present study aimed to explore how self-touch is perceived when the self is observed through other perspectives. It was found that self-touch in a mirror and third-person perspective lowers the sense of ownership of the observed hand, as well as the agency of the touch. Notably, the sense of ownership and agency were lower in the third-person perspective relative to mirror perspective. Further, no qualitative differences in touch perception were observed across the three perspectives. There was no relationship between the sense of ownership or agency and these qualitative differences. Lastly, there were no correlations between either of these experimental aspects (ownership, agency, qualitative) and autism traits and self-reported interoceptive abilities. The present study extends the knowledge and understanding of self-touch, and how visual perspective influences. The present study also highlights areas of applications, such as immersive virtual reality and clinical research.

Linda Handlin, Giovanni Novembre, Heléne Lindholm, Robin Kämpe, India Morrison — Human endogenous oxytocin and its neural correlates show adaptive responses to social touch based on recent social context

Both oxytocin (OT) and touch are key mediators of social attachment. In rodents, tactile stimulation elicits endogenous release of OT, potentially facilitating attachment and other forms of prosocial behavior, yet the relationship between endogenous OT and neural modulation remains unexplored in humans. Using serial sampling of plasma hormone levels during functional neuroimaging, we show that contextual circumstances of social touch facilitate or inhibit not only current hormonal and brain responses, but also calibrate later responses. Namely, touch from a romantic partner enhanced subsequent OT release for touch from an unfamiliar stranger, yet OT responses to partner touch were dampened following stranger touch. Hypothalamus and dorsal raphe activation reflected plasma OT changes during the initial interaction. In the subsequent social interaction, OT modulation depended on the previous interaction, mediated by precuneus and parietal-temporal cortex pathways, including a region of medial prefrontal cortex that also covaried with plasma cortisol. These findings demonstrate that hormonal neuromodulation during successive human social interactions is adaptive to social context, and they point to mechanisms that flexibly calibrate receptivity in social encounters.

Lina Koppel, Giovanni Novembre, Robin Kämpe, Mattias Savallampi, and India Morrison — Prediction and action in cortical pain processing

Predicting that a stimulus is painful facilitates action to avoid harm. But how distinct are the neural processes underlying the prediction of upcoming painful events, and those for taking action to avoid or prevent them? In this fMRI experiment, we addressed this by investigating brain activity as a function of current and predicted painful or nonpainful thermal stimulation, and the ability of voluntary action to affect the duration of the upcoming stimulation. Participants performed a task which involved the administration of a painful or nonpainful stimulus (S1), which predicted an immediately subsequent very painful or nonpainful stimulus (S2). On action-effective trials, pressing a button within the specified time window during S1 reduced the duration of the upcoming stimulation in S2. On action-ineffective trials, pressing the button had no effect on upcoming stimulation. Predicted pain increased activation in regions including anterior cingulate cortex (ACC), midcingulate cortex (MCC), and insula, and depended on whether a meaningful action was performed. Insula responses for predicted pain were also modulated by potential action consequences. Taken together, these findings suggest that cortical pain processing is not specifically tied to the sensory stimulus, but instead depends on the consequences of that stimulus for sensorimotor control of behavior.

Clara Larsson — Point of View: The Impact of Background Conditions on Distinguishability of Visualised Data in Detailed Virtual Environments

Data visualisation in a virtual environment (VE) opens up new ways of presenting and processing information. It makes it possible to explore data in an immersive way. However, it comes with challenges. One of these challenges is the perception and distinguishability of the data. The data needs to be distinguishable against the background, but in a VE where the user can move, interact and observe the data from different perspectives, the backdrop will be constantly changing.

The study used the in-development air quality data visualisation system CityAirSim and a self-completion questionnaire to answer the research question: What impact does a detailed virtual environment have on distinguishability of visualised data?

The collected data were not able to determine if one of the colourmap used (YellowRed, Rainbow) was more efficient. The darker environment was reported to make the visualisation easier to see in comparison to the lighter environment. The results also showed that different perspectives and their different backdrops impacted the distinguishability of the data, with participants answers being different depending on if the virtual camera was placed above or below the visualisation.

Zakarias Mortensen — Emotional Design in VR, A study on emotional design and coloursin a virtual reality setting

The purpose of this paper is to contribute with knowledge regarding how the emotional design element colours could be incorporated into the realm of Virtual Reality (VR) to achieve elicitation of positive emotions. A mixed methods approach was implemented in which respondents explored two virtual environments, one containing saturated colours and one containing greyscale colours. The overarching goal with the given approach was to look closer at how colours could be beneficial in the context of virtual environments (VE) and whether the already found benefits from implementation of emotional design on the web could be transferred to VR. The results indicate that saturated colours could evoke emotions in a virtual environment, but the outcome of which feelings the saturated colours could evoke was ambiguous. Two notions were found where the respondents either found the version containing the saturated colours to fill them with excitement, warmth and comfort, whilst the other notion proposed that the environment felt less realistic and did not fill its purpose of being a learning environment.

Patrick Oden & Linus Holm — Intrinsic Rewards for Problem Resolutions

Resolving abstract problems often appear intrinsically satisfying, but what determines the level of internal reward and why would the brain reward itself for solving problems without external utility? A native drive for solving problems might spur inventions and insights that eventually turn out to be adaptive to the individual as well as the group and the entire species. For such a drive to be adaptive in the long term, it seems metacognitive assessment of problem resolution needs to be accurate. We hypothesized that problem difficulty might drive resolution satisfaction in two online mathematical logical forced choice tasks involving 30 and 40 problems, and 96 and 102 participants, respectively. In both Experiments, we found that problem difficulty predicted resolution satisfaction, and that participants' resolution confidence ratings were closely related to resolution satisfaction.

Astrid Venell, Emma Olsson & Linus Holm — Raising interest by satisfying curiosity

Some little piece of information sometimes increases the appetite for more. Understanding what raises and maintains topical interest would be of great value in education, and would connect current research on curiosity with more obviously adaptive aspects of intrinsic cognitive motivation. We investigated what factors in a learning situation might spur further topical interest in an online test where 96 participants rated their experience of 100 two-alternative forced choice zoology trivia questions before and after receiving the correct answer. We found that pre-answer curiosity, post-answer surprise and satisfaction reliably increased the likelihood of requesting unrelated extra animal information at the cost of a time delay. Our findings suggest how covered knowledge gaps instead of saturating the desire for knowledge instead act to increase it.

Michael Wells & Linus Holm — Inherently satisfying retrieval favors reliability

It often seems satisfying to retrieve an item from memory, irrespective of the memory content. If the brain incentivizes retrieval attempts on the prospect of an internal retrieval reward, then that would constitute an intrinsic drive for reinforcing declarative memory access. But what determines the level of retrieval satisfaction? We tested the idea that retrieval attempt uncertainty drives retrieval satisfaction. For instance, the more distant the memory, the more satisfying should it be to successfully retrieve. Using five Swahili-English word pair encoding sessions spaced across one week in an online experiment, we tested 30 English-speaking participants' recall satisfaction and memory confidence in a final cued recall test. We hypothesized that retrieval satisfaction should increase with retrieval uncertainty as indicated by time since encoding, and encoding session retrieval performance. Instead, we found that retrieval satisfaction decreased with time since encoding and session retrieval performance. Moreover, we found that retrieval confidence and satisfaction ratings were essentially interchangeable in the experiment. Thus, the brain appears to reward retrieval reliability rather than fruitful retrieval effort.

Bibliography

Florian Fiebig, Pawel Herman, and Anders Lansner. An indexing theory for working memory based on fast hebbian plasticity. *eNeuro*, 7(2), 3 2020. ISSN 23732822. doi: 10.1523/ENEURO.0374-19.2020. URL https://pubmed.ncbi.nlm.nih.gov/32127347/.

Bertram Opitz. Context-dependent repetition effects on recognition memory. *Brain and Cognition*, 73(2):110–118, 7 2010. ISSN 02782626. doi: 10.1016/j.bandc.2010.04.003. URL https://pubmed.ncbi.nlm.nih.gov/20493623/.

Philip J. Tully, Matthias H. Hennig, and Anders Lansner. Synaptic and nonsynaptic plasticity approximating probabilistic inference. *Frontiers in Synaptic Neuroscience*, 0(APR):8, 2014. ISSN 1663-3563. doi: 10.3389/FNSYN.2014.00008.

E Tulving. Episodic and semantic memory. In E. Tulving and W. Donaldson, editors, *Organization of Memory*, pages 423–445. Academic Press., 1972. URL https://psycnet.apa.org/record/1973-08477-007.

Generative Models as Epistemic Artifacts

Raphael Aybar[1]

[1]*University of Vienna*

Predictive Processing (PP) proposes that the brain embodies a generative model of the environment that predicts internal or environmental states to reach adaptation. Critics accuse the neurocentric implications of PP as it sets the boundaries of cognition within the brain and the body while putting aside the roles of culture and environmental resources in extended cognition. I contest these critics arguing that generative models are epistemic artifacts for modeling cognitive systems or scientific models that create new models (Tee, 2020). Although proponents of PP certainly generative models as tools, they conflate them with mental kinds. The argument develops as follows: generative models are theoretical constructs in scientific modeling practices, which do not target any particular physical system. In other words, neither they have to be realized in cortical activity nor elsewhere. From an observer's perspective, they are architectures of information processing that can be realized in the brain or agent-environment coupled systems. When approached epistemic artifacts, their 'generativity' is not a capacity to represent but produce something new. This property may serve to model non-brain-centered cognitive systems, such as niche constructing adaptive systems (Bruineberg et al., 2018).

Modeling cognition can be a misleading epistemic practice: it aims to create bona fide representations of cognitive systems by capturing cognitive structures shared by humans, other animals, and other organisms. However, when scrutinizing in this practice, we can appreciate something else developing in it: the alleged cognitive structures are, in fact, observer-relative, created in scientific practices, where sometimes are postulated as non-observable entities. This is the starting point of an epistemology of cognitive sciences. Its main question concerns the epistemic gain of modeling practices in this field of research.

Cognitive scientists typically employ reductive strategies when modeling. Consider the practice of modeling perception. Perception can be the subject of multiple approaches (e.g., phenomenological, mechanistic, or psychological), but a perception model cannot. Of course, scientists can always employ multiple model strategies (see Weisberg, 2007). A modeler can define perception as a computational process by narrowing down the understanding of the phenomenon when doing this. The modeler can approach it instead from a first-person perspective, but perhaps at the price of overlooking its physiology (this is not meant to say that scientists with different views of perception cannot collaborate). This reductive strategy has a positive side: it makes available for the scientists a set of tools that can help her transform her intuitions into tangible and manipulable artifacts called scientific models. The epistemic gain of modeling practices lies precisely in the construction and manipulation of models (Knuuttila, 2011). In the rest of this talk, I will understand models as epistemic artifacts constructed and manipulated by scientists to solve questions or gain understanding. Their concrete and abstract properties make these surrogates apt to foster and limit scientific reasoning and imagining (Knuuttila, 2017). I will analyze recent predictive processing models of cognition based on this perspective.

Predictive Processing is a two-fold theoretical framework in the cognitive sciences: it is a general conceptualization of mental phenomena associated with a set of statistical and computational tools. Cognitive scientists use them to construct scientific models that help them investigate questions. These tools have mainly used Bayesian methods to solve probabilistic problems; the general conceptualization is an analogy between cognitive systems and probabilistic engines (Clark, 2016). Proponents of PP tend to depict cognitive agents as embodied systems that structure their information about the world by forming mental structures resembling the world. This idea is appealing: it is not sufficient for an adaptive engagement to just have a system of unorganized information. Information has to be arranged so that it becomes a tool for acting, planning, perceiving, and even thinking. In this context, PP proponents suggest that minds gather information probabilistically by creating and updating generative models.

Generative models are the core components of PP systems. They mediate between the organization of sensuous information and perception and action. Perception is the recognition of the causes of the sensorium. For example, when a person smells something, identifying the smelling object causing the smell is a perception. Action is a movement intended to transform environmental states into a set of expected or desired states (Friston, 2012). Generative models are architectures of information processing presenting plausible ways by which cognitive agents perceive and act. In other words, a predictive system generates predictions to identify the object causing sensations or predict that certain policies will lead to desired states. To do either of both, the cognitive agent needs relevant ecological information. If the agent does not know its environment, it may produce weak predictions and misleading policies. A generative model is, in PP, a model or mental structure that organizes this information. Thus, it serves to accomplish the function of a predictive engine that perceives and acts.

In the philosophical literature of the predictive mind, the elegant solution of generative models to model perception-action systems has been mistakenly construed as a mechanistic account of the brain (Badcock et al., 2019); as a way of naturalizing mental representations (Hohwy, 2013); or as a realist view of the probabilistic brain (in the so-called Bayesian brain hypothesis; see Colombo and Seriès, 2012). I do not have space to criticize either of these views here. However, their shared problem is that they overlook the role of generative models as scientific tools, treating them instead as mental kinds. I think this misunderstanding is the main reason for the criticisms PP has been receiving (see Kogo and Trengove, 2015).

Generative models are a type of scientific model, described by Tee (2020) as models that create new models. Tee treats them as scientific models, which contemporary philosophers of science have extensively investigated. Godfrey-Smith (2006) tells us that current scientific practices proceed by a strategy of indirect representation: to solve scientific questions, scientists do not directly approach phenomena but use surrogate models. Morgan and Morrison (1999) tell us that models are entities partially independent from theories and reality but mediate between them. Knuuttila (2017) treats them as epistemic artifacts constructed and manipulated by scientists in the light of pending questions. She argues that the epistemic gain of modeling is not delivered by representational capacity but by the construction and manipulation of models. In these practices, scientists exploit the affordances of these artifacts that shape their reasoning and imagining.

Tee (2020) follows this line of thought when describing the generativity of models as the capacity not to represent but to create something new. There are two crucial implications of this definition for PP generative models: 1. if we think of them as scientific artifacts, they are not representations of cognitive systems but productive devices that deliver understanding by their construction and manipulation. 2. If we instead think of them as mental kinds, neither are they internal mental representations but action-oriented or productive structures. In any case, as scientific constructs, they are not models of any natural system, but models for doing something (Fox Keller, 2000), as Tee (2020) explicates in the following: “What counts as a generative model is determined by the scientist's use of the model. [...] The process of abstraction and idealization which confers the generative capacity on a generative model is a two-staged process that is not a characteristic of non-generative models: (1) the omission of the irrelevant features of a target system is carried out in the direction of producing, rather than representing, a new model based on the existing model; (2) new idealized model elements which are relevant to the generated model are added. These new model elements are relevant to the objective of the model construction and transform a generative model into a new model.”

Tee is interested in accounting for the generativity of models from the point of view of scientific practices, understanding it as the capacity to produce new models through abstraction and idealization. As scientific models in PP, generative models are architectures of information processing, specifying the components and interactions of cognitive systems underlying perceptual inference and action. A prediction, following this description, is not a representation of the world but a productive or action-oriented guess made by a cognitive agent to engage in the world (Gallagher and Allen, 2018). Scientists explicate this cognitive behavior by postulating a non-observable process using the Bayesian statistical methods, various abstract formulations of adaptive behavior, and cortical activity data. Unfortunately, we cannot explore the interdisciplinary transfer of knowledge developing in PP in this talk.

So far, we have described generative models as scientific models that create new models by abstracting and idealizing, which are analogous to the activities of learning and optimizing priors in PP. Also, I pointed out that generative models, as abstract scientific models, are not bound to a particular cognitive system; in fact, they can be used for modeling non-cognitive systems. This is because their abstract formulation does not specify any material system realizing their behavior. To do so, scientists aggregate existing knowledge. For instance, Urgen and Saygin (2020) have used generative models to account for information processing architectures underlying the Action Observation Network.

Since there are not bound to any specific physical system, my point is that they can be used as well for modeling non-brain-centered cognitive systems. A recent example is the agent-based model of niche adaptation from Bruineberg et al. (2018). They designed an agent-based model and performed computer simulations of an agent-environment coupling. This model was used for simulating adaptive behavior in environments under the condition that actions transform environmental states (niche construction). An adaptive agent maintains itself in states that do not threaten its existence (viable states). In agent-based modeling, it becomes an agent whose goal is to reach desired states and form policies to reach them. The problem for this agent is that, by following the policies, it modifies the distribution of states in the environment. Generative models provide this agent with a probabilistic solution for estimating how its actions transform the environment.

In this talk, I argued that generative models are epistemic artifacts employed in cognitive science. They are architectures of information processing that provide probabilistic solutions of tasks where agents deal with uncertain environments and have limited information. Though this abstract formulation has been mainly used to model cortical activity, it can be employed to model extended cognitive systems. This perspective implies that we should be careful when approaching them as mental kinds, not only because we may disregard their roles as epistemic artifacts but because we might conflate the properties of scientific models and natural phenomena.

References

Badcock, P., Friston, K. & Ramstead, M. (2019). "The Hierarchically Mechanistic Mind: A Free-Energy Formulation of the Human Psyche." *Physics of Life Reviews* 31:104–21.

Bruineberg, J., Rietveld, E., Parr, T., van Maanen, L. & Friston, K. (2018). "Free-Energy Minimization in Joint Agent-Environment Systems: A Niche Construction Perspective." *Journal of Theoretical Biology* 455:161–78.

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. New York, NY: Oxford University Press.

Colombo, M. & Seriès, P. (2012). "Bayes in the Brain. On Bayesian Modelling in Neuroscience." *British Journal for the Philosophy of Science* 63(3): 697–723.

Fox Keller, E. (2000). "Models Of and Models For: Theory and Practice in Contemporary Biology." *Philosophy of Science* 67: S72–86.

Friston, K. (2012). "A Free Energy Principle for Biological Systems." *Entropy* 14(11): 2100–21.

Gallagher, S. & Allen, M. (2018). "Active Inference, Enactivism and the Hermeneutics of Social Cognition." *Synthese* 195(6): 2627–48.

Godfrey-Smith, P. (2006). "The Strategy of Model-Based Science." *Biology and Philosophy* 21(5): 725–40.

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Knuuttila, T. (2017). "Imagination Extended and Embedded: Artifactual versus Fictional Accounts of Models." *Synthese*. https://doi.org/10.1007/s11229-017-1545-2.

Knuuttila, T. (2011). "Modelling and Representing: An Artefactual Approach to Model-Based Representation." *Studies in History and Philosophy of Science Part A* 42(2): 262–71.

Kogo, N & Trengove, C. (2015). "Is Predictive Coding Theory Articulated Enough to Be Testable?" *Frontiers in Computational Neuroscience* 9(September).

Morgan, M. & Morrison, M. (1999). *Models As Mediators*. Cambridge: Cambridge University Press.

Tee, S. (2020). "Generative Models." *Erkenntnis*. https://doi.org/10.1007/s10670-020-00338-w.

Urgen, B. & Saygin, A. (2020). "Predictive Processing Account of Action Perception: Evidence from Effective Connectivity in the Action Observation Network." *Cortex* 128(July): 132–42.

Weisberg, M. (2007). "Three Kinds of Idealization." *Journal of Philosophy* 104(12): 639–59.

Basic Visual Reflexes for a Humanoid Robot

Christian Balkenius[1], Trond A. Tjøstheim[1], Amandus Krantz[1], Birger Johansson[1]

[1]*Lund University Cognitive Science*

christian.balkenius@lucs.lu.se

We report on our endeavours to build a complete reflex system for the humanoid robot Epi (Johansson, Tjøstheim & Balkenius, 2020, Fig. 1 left) using a bottom up approach. The aim is to reproduce as many human reflexes as possible given the physical and perceptual limitations of the robot. These reflexes make up a complex set of sensory-motor behaviors that constitute the basis for higher cognition and are fundamental for cognitive processes such as attention, object identification, eye-hand coordination, navigation and social interaction. The work presented here is thus a first step toward a more complete cognitive system.

Although we are not the first to implement these reflexes in a robot (see e.g. Brazeal et al., 2001), as far as we know this is the first attempt that is motivated by the architecture of the human nervous system, and our aim is broader than previous attempts in that we want to reproduce the detailed neural architecture involved in the processes.

The reflexes are implemented using neuromorphic techniques using the Ikaros system (Balkenius, Tjøstheim & Johansson, 2020), but some of the sensory processing uses a hybrid approach for computational efficiency. This allows all the reflex systems to run in real time on the humanoid robot. Here we focus mainly on visual reflexes and mention other reflexes only in passing when they are necessary to understand the operation the of the visual reflexes.

We have previously implemented a model of pupil control that is able to control the pupils of the robot and reproduces a large number of studies of pupil dilation (Johansson & Balkenius,2017; Tjøstheim, Johansson & Balkenius, 2019; Balkenius, Fawcett, Falck-Ytter, Gredebäck & Johansson, 2019). The current work is an extension of that model with additional reflexes.

There are two main sources of input to the different visual reflex systems in the robot. First, the robot has two eyes that can move laterally and include one camera each. Since the field of view is limited, we are also experimenting with other sensors that will be used in the future for peripheral "vision". Second, there is position feedback from the servos that is used for "proprioception".

The visual processing required for the reflexes are based on computing optic flow in the camera images (Fig. 1 right). Visual processing is done at different resolutions on inputs from an image pyramid (Adelson at al., 1984, blue area in Fig 1). Next, we detect interest points in the image using a FAST-detector (Rosten & Drummond, 2006; purple area in Fig. 1). All following calculations are done only at these interest points which substantially improve the speed of the system. The next step is the calculation of optic flow between the interest points in subsequent images (green area in Fig. 1). The result is a set of flow vectors expressed in degrees per second to make them invariant to the spatial and temporal resolution of the processing. The optic flow at the different scales are merged and sent to two further parallel processing stages. The first of these computes focal optic flow and is used for the optokinetic reflex (red area in Fig. 1). The second stage is a looming detector that tries to find object that are approaching the robot (yellow area in Fig. 1). This information is used for a visual avoidance reflex. Focal visual flow is calculated by first masking flow vectors in the periphery, and then clustering the remaining vectors using a mean shift operation (Fukunaga & Hostetler, 1975). Looming is computed using a Hough transform method (Hough, 1962) that estimates the center of looming objects from the intersections of pairs of flow vectors, followed by another mean shift stage that finds clusters of center points with the fastest looming speed.

The results of the sensory processing contribute to five different reflexes that have been implemented in the robot:

First, **the optokinetic reflex** (OKR) makes the eye move in the direction of focal flow. The eyes of the robot follow the motion for a short while before quickly returning to their original position. This is a form of visual nystagmus. The OKR is also the first step toward smooth pursuit eye movements and uses visual flow computed by the cortical area MT. This is further relayed through the nucleus of the optic tract and dorsal terminal nucleus before indirectly reaching the oculomotor nucleus (Distler and Hoffmann, 2011).

Second, the **optokinetic cervical reflex** (OKCR) makes the neck accompany the motion of the eye toward a visual target in order to make the head point in the direction of the stimulus. We also hypothesize a complementary mechanism that makes the neck move in the direction of the eyes when they are oriented away from the center line.

Third, the **cervico-ocular reflex** (COR) uses proprioception in the neck to move the eyes in the opposite direction. In humans, this reflex usually has a very minimal role since the vestibular ocular reflex is the main mechanism that compensate for head (and body) movements (Huygen, Verhagen, & Nicolasen, 1991). However, since Epi does not currently use any gyro or accelerometer, that would correspond to the vestibular system in humans, the COR serves to compensate for movements of the head when it, for example, is moved by hand or moves to align with the eyes.

Fourth, the **looming reflex** (LR) produces an avoidance response to approaching objects which causes the head to turn away from the stimulus as well as blinking. In the brain, the visual detection of a looming object is performed in the pretectal area while the superior colliculus controls the actual avoidance reaction. In Epi, avoidance consists in turning the head and blinking and is implemented by temporarily turning off the lights in the eyes.

Finally, the above visual reflexes work together with the **cervico-collic reflex** (CCR) which is the stretch reflex of the neck. This reflex aims at keeping the neck in the same position despite disturbances. The CCR is implemented in Epi using a PID-controller on the electrical current used by the servos to obtain compliant behavior of the neck. It can be pushed in different directions and will turn back in response. Together these reflexes form the basis for tactile interaction with the robot. The position feedback from the servo is also used for neck “proprioception”.

In the future, we want to do extensive experimentation with the reflex model both in simulations and in the robot to test how the implemented reflexes contribute to various physiological and cognitive phenomena. The aim is to reproduce both normal behavior and the effect of lesions and other conditions.

The reflex system will subsequently form the basis for computational models of higher level processes such as visual attention and memory. We also want to move the modelling from purely reactive behaviors, as described here, to deliberate processing that works on top of these basic visual reflexes

A gyro will also be included in the head of the robot which will make it possible to implement the vestibulo-ocular reflex (VOR) as well as the vestibulocollic reflex (VCR). In the mobile version of the robot we will also include information about locomotion to modulate the reflexes in various ways.

The source code for the reflexes described above are available for download as part of the Ikaros project (www.ikaros-project.org).

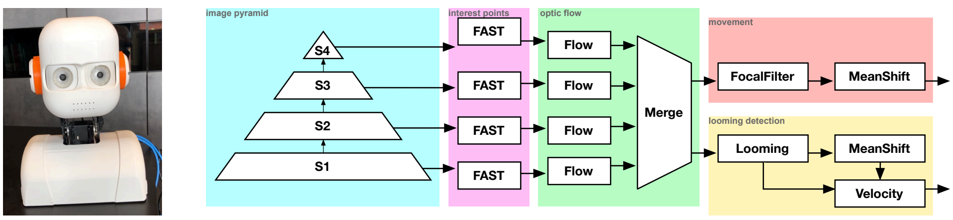


Figure 1. (left) The robot Epi. (right). The visual processing in the system (see text for explanation).

References

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., and Ogden, J. M. (1984). Pyramid methods in image processing. *RCA engineer*, 29(6), 33-41.

Balkenius, C., Fawcett, C., Falck-Ytter, T., Gredebäck, G. and Johansson, B. (2019). Pupillary Correlates of Emotion and Cognition: A Computational Model, *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, 903-907.

Balkenius, C., Tjøstheim, T., and Johansson, B., (2020). Ikaros: A Framework for Controlling Robots with System-Level Brain Models. *International Journal of Advanced Robotic Systems*, 17(3).

Breazeal, C., Edsinger, A., Fitzpatrick, P., & Scassellati, B. (2001). Active vision for sociable robots. IEEE Transactions on systems, man, and cybernetics-part A: Systems and Humans, 31(5), 443-453.

Distler, C., & Hoffmann, K. P. (2011). Visual pathway for the optokinetic reflex in infant macaque monkeys. *Journal of Neuroscience*, 31(48), 17659-17668.

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on information theory, 21(1), 32-40.

Hough, P.V.C., (1962), Method and Means for Recognizing Complex Patterns, U.S. Patent No. 3069654.

Huygen, P. L. M., Verhagen, W. I. M., & Nicolasen, M. G. M. (1991). Cervico-ocular reflex enhancement in labyrinthine-defective and normal subjects. *Experimental brain research*, 87(2), 457-464.

Johansson, B., Tjøstheim, T. and Balkenius, C. (2020). Epi: An Open Humanoid Platform for Developmental Robotics. *International Journal of Advanced Robotic Systems*, 17(2).

Johansson, B., Tjøstheim, T. A. and Balkenius, C. (2019). A Computational Model of the Acoustic Startle Reflex, *Proceedings of the 9th International IEEE EMBS Conference On Neural Engineering (NER'19).*

Johansson, B. and Balkenius, C. (2017). A Computational Model of Pupil Dilation. *Connection Science*.

Rosten, E. & Drummond, T (2006). Machine Learning for High-speed Corner Detection. *Computer Vision – ECCV 2006*. Lecture Notes in Computer Science. 3951. pp. 430–443.

Tjøstheim, T. A., and Johansson, B. and Balkenius, C. (2019). A Computational Model of Trust-, Pupil-, and Motivation Dynamics, *Proceedings of HAI 2019*.

Cognition as a Result of Information Processing in Living Agent's Morphology. Species-specific Cognition and Intelligence

Gordana Dodig-Crnkovic

Department of Computer Science and Engineering, Chalmers University of Technology

gordana.dodig-crnkovic@chalmers.se

Abstract. Cognitive science, according to *The Stanford Encyclopedia of Philosophy* (Thagard, 2014) is an interdisciplinary study of *human mind* and *intelligence*. It investigates *knowledge generation* in humans through perception, thinking (reasoning), memory, learning, and problem solving. Under this framing of cognitive science, variety of unsolved/unsolvable problems appear, as it ignores the role of the physical world and the body of a cognizing agent, neglecting the dynamical systems aspects, and emotions, as well as the phenomenon of distributed cognition. Moreover, it is ignoring the existing mathematical results which indicate that the human brain cannot be a classical computer (the Turing machine), with cognition modeled as computation over mental representations. On the other hand, radical biologism argues "Cognition = Life" (Maturana and Varela, 1980) (Stewart, 1996), thus the totality of biological processes. The functional connection is still missing between those two views, the high-level view of cognition with thoughts, mind, and intelligence (symbol processing) and the low-level view where each living organism (and all its building blocks, cells) is processing information that has a function for survival. Development of both cognitive science and related research fields of psychology, philosophy of mind, linguistics, neuroscience, bioinformatics, anthropology, and artificial intelligence go in the direction of embodied, embedded, extended, and enacted cognition (EEEE).

Moreover, implementations of cognition and intelligence in artifacts are contributing to the more detailed view of the relationship between cognition and its substrate. To achieve a connection between symbolic and sub-symbolic cognition, the concepts of computation and cognition must be generalized. Computation can be understood as process of morphogenesis (the development/generation of morphological characteristics, such as shape, form, and material composition in material bodies) on different levels of organization: physical, chemical, biological, cognitive, and virtual-machine-level computation built on top of them. Morphological computation/information processing approach to cognition provides a framework that connects low-level with high-level approaches to cognition and meets challenges and open questions listed by (Thagard, 2014).

In addition, the idea of cognition is generalized from the exclusively human capacity to the capacity of a variety of goal-directed adaptive self-reflective systems, from simplest living organisms (cells) humans (Dodig-Crnkovic, 2018). The aim is to better understand generative mechanisms of cognition in the light of evolution, as according to (Dobzhansky, 1973) "Nothing in biology makes sense except in the light of evolution."

(Thagard 2013; 2014) makes an extension of the idea of "thinking" to include emotional experience "The term cognition, as used by cognitive scientists, refers to many kinds of thinking, including those involved in perception, problem solving, learning, decision making, language use, and emotional experience." Similarly, (Lerner et al. 2015) argues for the importance of emotion in the decision making. The extension of "thinking" to emotional decision-making, also found in (Kahneman 2011), bridges some of the distance between cognition as (rational) thinking and its embodiment, but the basic problems remain of *generative mechanisms* that can dynamically connect body (matter) and mind. Thagard's description of cognitive science that includes emotions, does not make connections to biology, chemistry, (quantum- nano-, etc.) physics or chaos theory, self-organization, artificial life or data science, extended mind, distributed cognition, network science, sociology, or ecology, thus offering a rather high-level and thus necessarily simplified view.

Despite various attempts to bridge this body-mind gap made by (Clark 1997, 1989, 2013), (Scheutz, 2002) (Pfeifer and Iida, 2005), and several others, who have been offering a connection between sub-symbolic (signal processing) and symbolic (higher level) notions of cognition, this is still not reflected in the view of cognition found in major encyclopedias.

For the study of physical implementations of computation and control in robotic systems, the idea of morphological computation has been proposed in robotics (Paul, 2004) (Pfeifer and Bongard 2006) (Pfeifer, Iida, and Gomez 2006) (Pfeifer and Gomez 2009) (Hauser, Füchslin and Pfeifer, 2014) (Müller & Hoffmann, 2017a, 2017b) which defines computation process in a more general way than the conventional Turing machine symbol manipulation, considering physical embodiment of computational devices. Morphological computation in robotics uses the body to perform computation by its physical morphology (material/ shape/form) and thus replaces detailed central control, as described in (Matsushita et al. 2005) (Pfeifer, Iida, Lungarella, 2014).

In the earlier work of the author (Dodig-Crnkovic, 2012, 2014, 2016, 2017, 2018) arguments have been presented for the broadening of the definition of cognition, in the direction of EEEE cognition to include sub-symbolic processes that precede language in humans such as feelings, emotions, intuitions, and values (von Haugwitz and Dodig-Crnkovic, 2015). This more inclusive notion of cognition, related to its generative processes, evolution, and possible generalization to artificial cognitive systems provides a link between "cognitivist" (i.e. classical "computationalist") and EEEE approaches through *the idea of general morphological computation, that is computation performed by the morphology of an agent*. In this context morphological computation is info-computational self-organization "*in materio*" (Dale et al. 2017) of cognizing agents.

In living organisms, profound insights can be made by studying evolution and mechanisms of cognition at a variety of levels of organization, from single cells up to most complex living organisms. Cognition in nature appears throughout biological systems (Baluška and Levin 2016; Lyon 2005; Lyon 2015) and it is important to understand its evolutionary development from the basal/basic/elementary cognition on the cell level, to the human level (Manicka and Levin 2019; Levin et al. 2021).

Naturalized evolutionary approach to cognition is based on the view of hierarchical recursive structure of information processing in nature, in living organisms from cells, to tissues, organs, organisms, and their groups – all of them communicating at different levels of organization by exchanging specific types of information – physical (elementary particles, electro-magnetic), chemical (electric, molecular), biological, and symbolic (signs, languages).

By the ability to model cognition as embodied, embedded, enactive, and extended via interactions with the environment, morphological computing provides means of understanding how this capacity evolved, and how it develops during the life of an organism. Taking lessons learned from nature helps already in the engineering of artifactually cognitive and intelligent agents (Pfeifer, Iida, Lungarella, 2014).

In conclusion, current work in different fields informing cognitive science, from neuroscience (with neuroimaging) to robotics and AI, as well as novel insights about the inner working of cells, (especially neurons), and brains, all contribute to constructing a more complete picture of the physical basis of cognition and its underlying information processes. For the future, it remains to work out the details of this emerging view of cognition where not only thoughts (symbolic computation) but also perceptions, sensations, feelings, and emotions (sub-symbolic, physical computation "*in materio*", connecting cognizing agents with their bodies and the world will be recognized as a substantial part of cognition (Feldman Barrett, 2018) (Damasio, 2008). Tracing back cognition to its most rudimentary forms in unicellular organisms and carefully studying its complexification on the evolutionary scale helps understand physical basis of cognition and the emergence of symbolic computation that from chemical language of bacteria, leads to human-level linguistic competences.

Information processing in cognizing agents can be modeled as substrate-specific/species-specific morphological computation, from a single cell and up. If we want to learn how cognition functions in humans as the most complex living organisms, it is instructive to see how this ability developed through evolution, resulting in a variety of cognitive architectures of organisms from bacteria to humans (Ginsburg and Jablonka 2019)(Manicka and Levin 2019).

References

Baluška, F. and M. Levin. (2016) "On Having No Head: Cognition throughout Biological Systems." Frontiers in Psychology 7: 902.

Clark A. (1989) Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing. Cambridge, MA: MIT Press.

Clark A. (1997) Being There: putting brain, body and world together again. Oxford University Press.

Clark A. (2013) "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." The Behavioral and Brain Sciences 36 (3) (June 10): 181–204.

Dale, Matthew, Susan Stepney, Julian Miller, and Martin Trefzer. 2017. *Reservoir Computing in Materio: A Computational Framework for in Materio Computing*.

Ginsburg, Simona, and Eva Jablonka. 2019. *The Evolution of the Sensitive Soul*. MIT Press, Cambridge, MA,.

https://doi.org/10.7551/mitpress/11006.001.0001.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Macmillan.

Lerner, Jennifer S, Ye Li, Piercarlo Valdesolo, and Karim S Kassam. 2015. "Emotion and Decision Making." *Annual Review of Psychology* 66 (1): 799–823. https://doi.org/10.1146/annurev-psych-010213-115043.

Levin, Michael, Fred Keijzer, Pamela Lyon, and Detlev Arendt. 2021. "Uncovering Cognitive Similarities and Differences, Conservation and Innovation." *Phil. Trans. R. Soc. B* 376: 20200458.

Manicka, Santosh, and Michael Levin. 2019. "The Cognitive Lens: A Primer on Conceptual Tools for Analysing Information Processing in Developmental and Regenerative Morphogenesis." *Philosophical Transactions of the Royal Society B* 374 (1774).

Pfeifer, Rolf, and Josh Bongard. 2006. *How the Body Shapes the Way We Think – A New View of Intelligence*. MIT Press.

Pfeifer, Rolf, and Gabriel Gomez. 2009. "Morphological Computation - Connecting Brain, Body, and Environment." In *Creating Brain-like Intelligence: From Basic Principles to Complex Intelligent Systems*, edited by K. B. Sendhoff, O. Sporns, E. Körner, H. Ritter, & K. Doya, 66–83. Berlin: Springer.

Pfeifer, Rolf, Fumia Iida, and Gabriel Gomez. 2006. "Morphological Computation for Adaptive Behavior and Cognition." *International Congress Series*, no. 1291: 22–29.

Thagard, Paul. 2013. "Cognitive Science." In *Encyclopædia Britannica*. https://www.britannica.com/science/cognitive-science.

———. 2014. "Cognitive Science." In *The Stanford Encyclopedia of Philosophy*, Edward N. https://plato.stanford.edu/archives/fall2014/entries/cognitive-science/.

Damasio, A. (2008) Descartes' Error. Emotion, Reason and the Human Brain, Random House.

Dobzhansky, T. (1973), Nothing in Biology Makes Sense Except in the Light of Evolution, American Biology Teacher, 35 (3): 125–129

Dodig-Crnkovic G. (2012) The Info-computational Nature of Morphological Computing, in Müller V. C. (ed.), Theory and Philosophy of Artificial Intelligence (SAPERE; Berlin: Springer), pp. 59-68.

Dodig-Crnkovic G. (2014) Modeling Life as Cognitive Info-Computation, In: Computability in Europe 2014, Arnold Beckmann, Erzsébet Csuhaj-Varjú and Klaus Meer (Eds.) Proceedings of the 10th Computability in Europe 2014, Language, Life, Limits, Budapest, Hungary, June 23 - 27, 2014, LNCS, Springer

Dodig-Crnkovic G. (2016) Information, Computation, Cognition. Agency-Based Hierarchies of Levels. FUNDAMENTAL ISSUES OF ARTIFICIAL INTELLIGENCE, Müller V. C. (ed.), Synthese Library 377, pp 139-159. Springer International Publishing Switzerland 2016, DOI 10.1007/978-3-319-26485-1_10 http://arxiv.org/abs/1311.0413

Dodig-Crnkovic G. (2017) Nature as a Network of Morphological Infocomputational Processes for Cognitive Agents, The European Physical Journal Special Topics, DOI: 10.1140/epjst/e2016-60362-9 Eur. Phys. J. 2017, 226, 181–195.

Dodig Crnkovic, G. (2018) Cognition as Embodied Morphological Computation. Philosophy and Theory of Artificial Intelligence 2017: 19-23. http://dx.doi.org/10.1007/978-3-319-96448-5

Feldman Barrett, L. (2018) How Emotions are Made: The Secret Life of the Brain, Pan Books

Ginsburg, S. and E. Jablonka (2019) The Evolution of the Sensitive Soul. MIT Press, Cambridge, MA.

Hauser H., Füchslin R. M., Pfeifer R., Eds. (2014) Opinions and Outlooks on Morphological Computation, ISBN (Electronic) 978-3-033-04515-6 http://www.merlin.uzh.ch/contributionDocument/download/7499

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Macmillan.

Lerner, Jennifer S, Ye Li, Piercarlo Valdesolo, and Karim S Kassam. 2015. “Emotion and Decision Making.” *Annual Review of Psychology* 66 (1): 799–823. https://doi.org/10.1146/annurev-psych-010213-115043.

Levin, Michael, Fred Keijzer, Pamela Lyon, and Detlev Arendt. 2021. “Uncovering Cognitive Similarities and Differences, Conservation and Innovation.” *Phil. Trans. R. Soc. B* 376: 20200458.

Lyon, P. (2005) The Biogenic Approach to Cognition. Cognitive Processing 7: 11–29.

———. 2015. The Cognitive Cell: Bacterial Behaviour Reconsidered. Frontiers in Microbiology 6: 264.

Manicka, S. and M. Levin (2019) The Cognitive Lens: A Primer on Conceptual Tools for Analysing Information Processing in Developmental and Regenerative Morphogenesis. Philosophical Transactions of the Royal Society B 374 (1774).

Matsushita K., Lungarella M., Paul C., Yokoi H. (2005) Locomoting with Less Computation but More Morphology, Proc. 2005 IEEE Int. Conf. on Robotics and Automation, pp 2008-2013

Maturana H. R. & Varela F. J. (1980) Autopoiesis and Cognition - The Realization of the Living. R. S. Cohen and M. W. Wartofsky (Eds.), Boston Studies in the Philosophy of Science 42. Dordrecht, The Netherlands: D. Reidel Publishing.

Müller V. C. & Hoffmann M. (2017a), 'What is morphological computation? On how the body contributes to cognition and control', Artificial Life 23 (1), 1-24

Müller, V. C. & Hoffmann, M. (2017b), Simple or Complex Bodies? Trade-offs in Exploiting Body Morphology for Control. In: Representation and Reality in Humans, Other Living Organisms and Intelligent Machines. Dodig-Crnkovic, G., Giovagnoli, R. (Eds.), Springer International Publishing, DOI 10.1007/978-3-319-43784-2 pp. 335-347

Paul C. (2004) Morphology and Computation, Proceedings of the International Conference on the Simulation of Adaptive Behaviour Los Angeles, CA, USA, pp 33–38

Pfeifer R. and Iida F. (2005) Morphological computation: Connecting body, brain and environment. Japanese Scientific Monthly, Vol. 58, No. 2, 48–54

Pfeifer, Rolf, and Josh Bongard. 2006. *How the Body Shapes the Way We Think – A New View of Intelligence*. MIT Press.

Pfeifer, Rolf, and Gabriel Gomez. 2009. “Morphological Computation - Connecting Brain, Body, and Environment.” In *Creating Brain-like Intelligence: From Basic Principles to Complex Intelligent Systems*, edited by K. B. Sendhoff, O. Sporns, E. Körner, H. Ritter, & K. Doya, 66–83. Berlin: Springer.

Pfeifer, Rolf, Fumia Iida, and Gabriel Gomez. 2006. “Morphological Computation for Adaptive Behavior and Cognition.” *International Congress Series*, no. 1291: 22–29.

Pfeifer R., Iida F. and Lungarella M. (2014) Cognition from the bottom up: on biological inspiration, body morphology, and soft materials Trends in Cognitive Sciences August 2014, Vol. 18, No. 8

Scheutz M. (Ed.) (2002) Computationalism: New directions. Cambridge: Cambridge University Press.

Stewart J. (1996) Cognition = Life: Implications for higher-level cognition. Behavioural Processes 35: 311-326.

Thagard, P. (2013) Cognitive science. Encyclopædia Britannica. September 30, 2013 URL = https://www.britannica.com/science/cognitive-science

Thagard P. (2014) "Cognitive Science", The Stanford Encyclopedia of Philosophy (Fall 2014 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2014/entries/cognitive-science/>.

von Haugwitz R. and Dodig-Crnkovic G. (2015) Probabilistic Computation and Emotion as Self-regulation ECSA 2015 ASDS Workshop. In Proceedings of the 2015 European Conference on Software Architecture Workshops (ECSAW '15). ACM, New York, NY, USA, DOI=10.1145/2797433.2797442

Gaze tracking partisans: Trust and individual conservatism mediate visual attention to search engine results pages

Ekström, A. G.,[1, 2, 3] Niehorster, D. C.,[2, 4] & Olsson, E. J.,[5]

[1]*Lund University Cognitive Science.* [2]*Department of Psychology, Lund University.*
[3]*KTH Speech, Music and Hearing.* [4]*Lund University Humanities Lab.*
[5]*Department of Philosophy, Lund University.*

axeleks@kth.se

Introduction

Research in social psychology and communication presents two disparate and seemingly incompatible accounts of biased information processing. In the *epistemic vigilance* framework (Sperber et al., 2010), people spend more attention on disparate (i.e., incongruent) information because of its heavier intrinsic processing load, whereas according to research on *selective attention* (Klapper, 1960), where an applied premium of relevance on the information at hand acts as a buffer against disagreeable information, the opposite is expected. In ongoing debates, this question has become increasingly relevant, with debate concerning possible polarizing effects of internet use, such that individuals may be driven to seek out information that confirms prior beliefs (e.g., Yom-Tov et al., 2014). In this work, we report an empirical test of the above-mentioned theoretical discrepancy in psychology research, using an experimental setting with bearing on the topic of biased internet use.

In research on the use of search engines (e.g., Google Search), a well-replicated finding is the tendency for users to select top-presented links in search result pages rather than links with lower positions, which has been dubbed the "top-link heuristic" (Salméron et al., 2013). Before information can be selected or not selected, however, it must be attended, and processed cognitively. Accordingly, eye-tracking technology, which allows for the study of unconscious behavioral processes that mediate conscious cognition and decision making (e.g., Pärnamets et al., 2015) has become a popular tool in this line of research. Of particular interest to online information search, Hotchkiss et al., (2005) found that when observing search results returned by a search engine, users allocated the most attention (operationalized as total viewing time) to the first pair of search results, with a steep drop-off for each subsequent link – an attention correlate of the top-link heuristic.

Outside of general search enquires, however, search engines are also used to search for information about divisive social issues. While much research has centered on eye movements in web usability research, psychological factors involved in attending to controversial search results remain largely unexplored. This includes any potential difference between liberals (i.e., left-wing partisans) and conservatives (i.e., right-wing partisans) to selectively attend to or ignore politically (in)congruent material. Here, we investigate the possibility that users in an online information search scenario afford greater attention, operationalized as total viewing time, to links from familiar or trusted sources and to links that support their own beliefs, compared to less trusted and less politically aligned links. To the authors' knowledge, the present study is the first to investigate this relationship.

Methods

Pre-study

Links were sampled from Google Search results, representing various online news sources, such as online magazines and newspapers, to represent a diversity of views across the political spectrum. 8 political issues, derived from Everett's (2013) Social and economic conservatism scale (SECS) were chosen as stimuli: (1) Abortion; (2) Benefits; (3) Climate change; (4) Equality between the sexes; (5) Nuclear family; (6) Immigration; (7) Religion; and (8) Taxation. For each topic, 10 links were sampled, for a total of 80 links. To judge the perceived ideology of each link, online participants (n = 3) were recruited for an online pre-study. Each link was presented once, along with a seven-point Likert scale, and participants were asked to rate how they perceived the ideological skew of each link from left to right, where lower ratings signified more liberal content, and higher ratings signified more conservative content. The ordering of the links was fully randomized for each participant. The pre-study took on average ~5 minutes to complete and participants were not compensated. For the ratings, Cronbach's alpha for the 80 items was deemed sufficient (α = .79) to proceed with analyses. In subsequent categorization of the links, scores for each link were averaged, and the average of the scores assigned to each link.

Stimuli

For the sake of environmental validity, it was determined that stimuli should be as visually similar as possible to real-life search engine result pages (SERPs). Thus, from the original 80, sets of six links per topic were selected for the study proper for a total of 64. This number was selected because previous research has suggested that attention to links past the fifth and sixth results in SERPs are rarely attended by searchers (e.g., Hotchkiss et al., 2005). Chosen links were selected to provide the widest possible range of political views in each presented stimulus. Further, to control for ordering effects, in each set of search results, the links were ordered according to a 6x6 Latin square, resulting in a total of six stimuli for each topic, from which one was sampled and presented to any given participant. This also required the number of participants to be a factor of six. In all, 8 stimuli were presented to each participant. All links and all linked-to sites were in Swedish.

Participants

54 participants were recruited from a combination of politically partisan organizations (e.g., political youth groups), college courses deemed likely to attract politically active students (e.g., political science), and online advertisements. Six participants were excluded because of significant data loss during recording, resulting in a sample of 48 participants (22 women) aged 18 – 53 (M = 23.19, SD = 5.79). All participants were native Swedish speakers.

Procedure

Prior to the experiment, informed consent was obtained. Participants (N = 48) were informed that their data would be anonymized, that they would not be exposed to harm of any kind, and that they had the right to opt out of participation at any point during the experiment. Participants were invited to sit down in front of the eye tracker (Tobii Pro Spectrum) and asked to use a provided chin rest. Participants were instructed to keep their head still for the duration of the experiment. The eye tracker was then calibrated. Further instructions were provided on screen.

Participants were instructed that they were about to take part in an experiment about searching for information in Google Search. They were to examine the provided sets of search results and select (by clicking) the link they would be most likely to select in natural search settings. In each set of search results, six links were visible. For each trial. After a participant had clicked on a link, the experiment moved on to the next trial. Once all 8 stimuli had been presented, participants were asked to fill in a survey. They provided on 7-point Likert scales their affect towards a set of political topics and phrases (derived from the Social and Economic Conservatism Scale; see Everett, 2013). Finally, they were asked to state the degree to which they trusted each source in a list– some but not all of which were included in the study proper. After the experiment, participants were debriefed as to the true motivation of the study and offered compensation. Each experiment lasted roughly ~30 minutes. Mouse clicks were also recorded but will be reported elsewhere.

Results

Zeroes (absence of any fixations) were observed for links presented in the fifth or sixth link positions exclusively, and thus taken as an artefact of link position. For further analysis, zeroes were therefore excluded, such that all included fixation data represented links that had been fixated on. Further, because fixation duration was highly variable between participants (see Holmqvist et al., 2011), data distribution was significantly skewed. Therefore, data were log10 transformed prior to analysis. A multiple linear regression, excluding zeroes, was then calculated to predict total fixation duration based on participants' Trust in the link source; Participant conservatism score; Link conservatism score, and an interaction between Participant conservatism score and Link conservatism score. The model met homolinearity, homoscedacity and multicollinearity assumptions; however, residuals were found to be not normally distributed. One data point was identified via Cook's distance as highly influential and removed prior to subsequent analysis. A statistically significant regression equation was found ($F_{6,\ 2152}$ = 36, p < .001), with

Table 1. Multiple linear regression analysis of relative total fixation duration on links

	Estimate	Std. error	*t*-value	*P*(z)
(Intercept)	3.18	.1	31.13	<.001
Trust in link source	.03	.004	7.84	<.001
Participant conservatism	-.04	.01	-3.31	<.001
Link conservatism	<.001	.01	-.98	.325
Participant conservatism * Link conservatism	.01	.003	2.42	.016
Link position	-.04	.004	-9.61	<.001
Nr. of characters in link	-.002	<.001	5	<.001

Adjusted R^2: .09

$R^2 = .09$ (see Table 1). Mean fixation duration was found to be longer for links associated with more trusted sources, compared with those associated with less trusted sources ($p < .001$). In addition, a significant interaction effect between Participant conservatism scores and Link conservatism scores was observed ($p = .016$). A significant main effect of Participant conservatism was observed, such that participants scoring high on conservatism tended to spend less time attending to presented links ($p < .001$) than those with lower conservatism scores. There was no independent effect of Link conservatism score. Finally, Link position ($p < .001$) and the Number of characters in each link ($p < .001$) were also included in the model as covariates and found to be highly significant.

Discussion

This work, aligning with the *selective attention* framework (Klapper, 1960), provides evidence that trust in a source of information serves as an important positive predictor of visual attention afforded to online sources of information. In research on message reliability, people's evaluations of communicated information are contingent on relevance and plausibility. We argue that participants' ratings of trust in sources of the presented links tapped into values of apparent plausibility, such that a more trusted source should on average be expected to provide more plausible and reliable news coverage and opinion making as judged by the participants. To our knowledge, this represents the first time such results have been demonstrated in experimental settings emulating online information search.

Further, this work adds to an understanding of how politically partisan users of search engines engage with political content in SERPs. We found that political conservatism negatively predicts visual attention to sources. While our results indicate as much, the literature is currently undecided as to whether political liberals and conservatives exhibit consistent differences in information processing (but see e.g., Jost & Amadio, 2012). Of potentially greater significance, we also observed an interaction between participants' Conservatism score and Link conservatism score, such that participants tended to attend search results that appeared to align with prior beliefs, over those that appeared not to. This is also in alignment with previous research on this topic (e.g., Knoblock-Westerwick et al., 2015).

The present paper helps direct important future work on topics of internet polarization. In real-life settings, users themselves may play the greatest role (by choosing their own search terms) in deciding which information they are presented with. An intriguing topic for future research thus concerns the extent to which partisans use biased search terms in information search. We propose that partisans may exhibit stages of filtering, such that undesirable or incongruent information is excluded at several stages in information gathering. We further suggest, as an indicator for future work, that in information search where topics are divisive, a final selection filter – that of link selection – may be largely independent of link presentation order.

Acknowledgements

The present work took place in the context of the research project *Filter Bubbles and Ideological Segregation Online: Do We Need Regulation of Search Engines?*, funded by Riksbankens Jubileumsfond for the Advancement of the Humanities and Social Sciences (2019-2021, Olsson PI).

The authors also gratefully acknowledge the Lund University Humanities Lab for providing access to lab equipment and valuable expertise.

References

Everett, J. A. (2013). The 12 item social and economic conservatism scale (SECS). *PloS one*, *8*(12), e82131.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hotchkiss G., Alston S., & Edwards G. (2005). Google Eye Tracking Report: How Searchers See and Click on Google Search Results (Enquiro Search Solutions). Accessed January 24, 2021, https://searchengineland .com/figz/wp-content/seloads/2007/09/hotchkiss-eye-tracking -2005.pdf.

Jost, J. T., & Amodio, D. M. (2012). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion*, *36*(1), 55–64.

Klapper, J. T. (1960). The effects of mass communication. Free Press.

Knobloch-Westerwick, S., Johnson, B. K., & Westerwick, A. (2015). Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, *20*(2), 171–187.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences*, *112*(13), 4170–4175.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*(4), 359–393.

Yom-Tov, E., Dumais, S., & Guo, Q. (2014). Promoting civil discourse through search engine diversity. *Social Science Computer Review*, *32*(2), 145–154.

Biological Motion Indicators for the Detection of Cyclists at Night

Paul Hemeren[1], Mikael Johannesson[1], Mikael Lebram[1], Fredik Eriksson[1]

[1]*University of Skövde*

paul.hemeren@his.se

1 INTRODUCTION

The use of visual aids can increase the ability of drivers to detect cyclists at night and reduce the seriousness of injuries if a crash occurs (Kwan & Mapstone, 2004; Wood, 2020). The use of reflectors placed on critical parts of the human body have been shown to increase cyclist conspicuity at night (Wood, Tyrell, Marszalek, et al., 2011). Drivers detect cyclists with reflective clothing that enhances the movement of the human body (biomotion) at considerably longer distances than a reflective vest, which is a very often used piece of clothing by cyclists who want to be detected in darkness. Recent research (Wood, Tyrell, Lacherez & Black, 2017) has also shown that driver eye movements are quicker to fixate on cyclists who are wearing the biomotion reflector clothing than the reflective vest.

The explanation of the effect of the biomotion-patterned reflector placement stems from the sensitivity of human vision to the movement of other humans. This has been demonstrated in a wealth of research, which was promulgated by Gunnar Johansson (Johansson, 1973). Human vision can detect the movement and portrayed actions as points of light attached to the joints of a human body in a little as 300 milliseconds. This sensitivity can be exploited by placing reflective material on the joints of human body so that visual conspicuity is increased for cyclists and other vulnerable road users. Edwaard, et al., (2020) showed clear results of the conspicuity benefits of biological motion when cyclists were pedalling during the daytime.

In our research reported here, we demonstrate the effectiveness of cyclists wearing different patterns of reflectors on the distance it takes drivers to detect cyclists in different visual conditions that occur in an existing city environment. This research was carried out in a driving simulator in order to include more naturalistic testing conditions and to achieve a relatively high level of experimental control. As a complement to our previous research (Hemeren, et al., 2014) we aimed to determine the distance at which drivers would detect cyclists dressed in different patterns of reflective clothing, i.e., biomotion, standard vest, and no reflective material at all on the cyclist, which is the minimum legal requirement.

2 METHODS

2.1 Participants

Twenty-four participants (19 males, mean age 29) were recruited from the student population at the University of Skövde and from the circle of acquaintances of the experimenters. All participants had a valid driving license, but driving experience/frequency varied, from once a month to daily. All participants signed consent forms, and the experiment was conducted according to Swedish law and ethical guidelines.

2.2 Design and Procedure

Based on our previous research, three clothing patterns were used (Fig. 1): biomotion, vest and the minimum legal requirement (legal), in which no reflector material was worn by the cyclists. The reflective material used in the biomotion and vest conditions was 3M™ Scotchlite™ Reflective Material 8910 Silver Fabric. The vest was a common item purchased at a local store. In order to maintain focus on placement and not on other material, the reflective material on the vest was replaced with the same reflective material used for the biomotion clothing. Importantly for comparison, the amount of visible reflective material in these two conditions was approximately the same. The reflective strips on the vest were much wider than the strips in the biomotion condition.

A 4.6 kilometer long route which included central areas of the city of Skövde was selected for the video recording of cyclists dressed in the different reflective clothing conditions. Twelve positions were also selected along the route for placement of cyclists. The positions were selected to include well-lit areas with street lights and other light sources for commercial areas, no street lights and positions where the visibility of cyclists was partially obstructed by making a turn or by bushes along the road. The bicycles were placed in stationary training stands so that distance measures could be reliably made while cyclists pedalled. Each clothing condition was video recorded at each position for a total of thirty-six conditions. The spacing of the positions along the route was unevenly spaced in order to reduce expectancy effects about the presence of a cyclist.

A GoPro video camera was placed on the windshield of a car, and all forty-eight conditions (including the no-cyclist present condition) were recorded at a speed of forty km/h. The cyclists were recorded as the car approached from behind. Only the backside of the cyclists was visible. The car headlights were set to bright to compensate for the reduced visual quality of the video recordings. The recordings were then edited to produce driving routes that contained counter-balanced orders of clothing condition and route position.

The experiment was conducted in the driving simulator (custom built) at the University of Skövde. The participants were instructed to drive the car and maintain a speed of 50 km/h as the primary task. The gas pedal of the car was used to increase the speed of the recorded film to give the impression of acceleration. The steering wheel could also be used to create changes in the image on the screen to create a feeling of control similar to a

Figure 1: Reflective clothing patterns from left to right: biomotion, vest and legal minimum.

fully functional simulator environment. The secondary task was to honk the horn as soon as a cyclist was detected.

Each participant drove the route two times, which resulted in viewing each clothing condition six times for a total of 18 observations per participant. Since it was not practically feasible for each participant to view each clothing condition at each position, because of likely practice effects, the conditions were evenly divided among two groups of participants. The starting point along the route for each participant was randomly determined.

3 RESULTS

There were twelve possible correctly detected cyclists for each of the thirty-six conditions due to the necessary balancing of conditions and to reduce practice effects. The detection results for each condition showed that drivers were generally accurate at detecting cyclists with the exception of two positions (9 and 11) along the route where detection accuracy for the legal condition (no reflective material) was 0 and 3 out of twelve. The detection accuracy for the vest condition for these two positions was also low, 4 and 6 respectively, while the biomotion condition resulted in perfect detection in all twelve trials.

The data for the distance at which cyclists were detected were entered into a between-groups ANOVA with position as a random factor. Given the design of the experiment, this was the most conservative statistical analysis compared to a repeated measures ANOVA. The means for detection distance as a function of clothing condition and route position are presented in Figure 2, which shows in positions 1, 3, 4, 5, 6, 7, 8, and 10 that the biomotion patterns triggered detection at much greater distances that the legal and vest conditions. In these positions, the street paths were fairly straight and could give rise to early detection. For positions 2 and 12, there were curves where there was no sufficient distance between the car and the bicycle in order to detect biomotion patterns at a greater distance.

The pattern of results across the different positions/places shows that the biomotion pattern of reflector placement ($\bar{x}$ = 60.28, *SEM*=1.27) is detected at much greater distances than legal ($\bar{x}$ = 31.67, *SEM*=1.32) or vest ($\bar{x}$ = 33.56, *SEM*=1.32). The exceptions to this occur when the visibility of cyclists is obstructed. The main effect of reflector placement was significant, $F(2,18) = 16.91$, $\eta 2 = 0.65$, $p < 0.001$. Pairwise comparisons between the conditions for reflector placement show significant differences between the legal and biomotion conditions,

$t(227) = 10.46$, $p < 0.001$, and between the vest and biomotion conditions, $t(227) = 9.62$, $p < 0.001$. The main effect of position was also significant, $F(9,18) = 2.69$, $\eta 2 = 0.57$, $p < 0.001$,. The interaction effect showed that the effect of the different reflector conditions varied as a function of position along the route, $F(18, 309) = 9.16$, $\eta 2 = 0.35$, $p < 0.001$.

4 CONCLUSIONS

The clothing that supports biomotion perception is superior to legal and vest. The surprising result here is that lack of any significant difference between legal and vest. This indicates that cyclists should wear reflective clothing that can trigger driver attention by the placing reflective material on the joints of the human body.

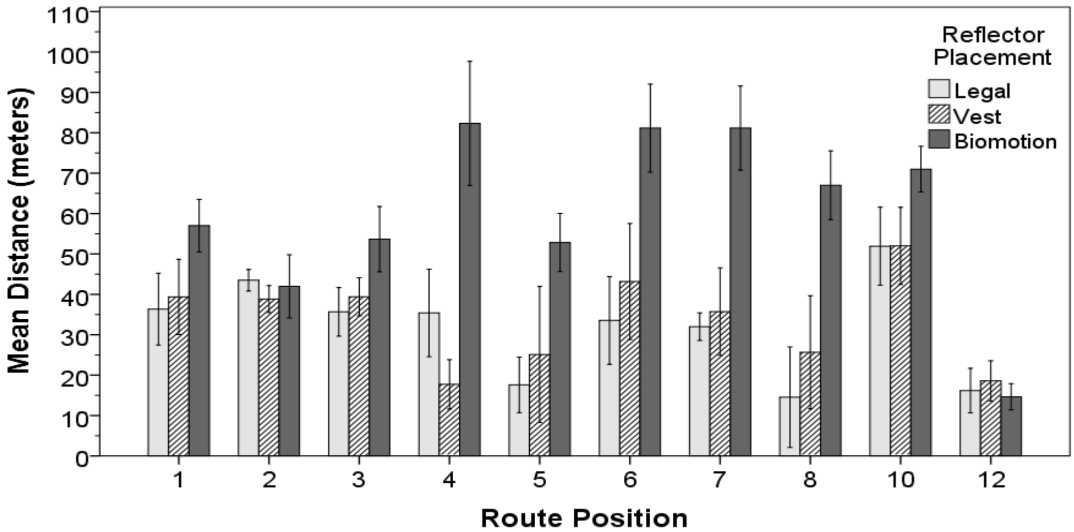


Figure 2: Distance detection in meters as function of reflector condition and place on route. Positions 9 and 11 are not included due to the very low detection accuracy for the legal and vest conditions.

References

Edewaard, D. E., Fekety, D. K., Szubski, E. C., & Tyrrell, R. A. (2020). Highlighting bicyclist biological motion enhances their conspicuity in daylight. *Accident Analysis & Prevention*, 142, 105575.

Hemeren, P. E., Johannesson, M., Lebram, M., Eriksson, F., Ekman, K., & Veto, P. (2014). The use of visual cues to determine the intent of cyclists in traffic. In 2014 *IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support* (CogSIMA) (pp. 47-51). IEEE.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, 14(2), 201-211.

Kwan, I. & Mapstone, J. (2004). Visibility aids for pedestrians and cyclists: a systematic review of randomised controlled trials. *Accident Analysis and Prevention*, 36, pp. 305-312.

Wood, J. M. (2020). Nighttime driving: visual, lighting and visibility challenges. *Ophthalmic and physiological optics*, 40(2), 187-201.

Wood, J. M., Tyrrell, R.A, Marszalek, R. P., Lacherez, P. F., Carberry, T. P. & Chu, B. S. (2011). Using reflective clothing to enhance the conspicuity of bicyclists at night. *Accident Analysis and Prevention*, 45, pp. 726-730.

Wood, J. M., Tyrrell, R.A, Lacherez, P. & Black, A. A. (2017). Night-time pedestrian conspicuity: effects of clothing on drivers' eye movements. *Ophthalmic & Physiological Optics*, 37, pp. 184-190.

Traveling Through the Dark: Using an interdisciplinary theatre and cognitive science approach to identify design strategies for human-machine shared experience in a self-driving car

Maurice Lamb[1], Rebecca Rouse[1]
[1]*Högskolan i Skövde, Institutionen för informationsteknologi*
Maurice.Lamb@his.se

Traveling through the dark I found a deer
dead on the edge of the Wilson River road.
It is usually best to roll them into the canyon:
that road is narrow; to swerve might make more dead.
-Excerpt from poem *"Traveling through the Dark"* by William Stafford (1998)

Recent research on human-machine interaction (HMI) across a range of fields, including both cognitive science and theatre, has stressed the need to re-frame such interactions as relational and based in shared experience (Gaggioli et al., 2021; Sciutti et al., 2018). In this case, the machine, whether software or hardware based, is characterized as an interaction partner instead of a tool. Reconceiving HMI as involving reciprocity and shared experiences moves away from transactional or one-sided models of interaction and requires exploring what can be meant by reciprocation and shared experience with a non-human partner. In particular, the concept of shared experience in HMI has been relatively under-explored due to both the typical framing of trust in HMI research and technological limitations of HMI systems. Refocusing the design of HMI systems on the ethos of shared experience can be supported by interdisciplinary research with theater. Anthropology and theater scholars have identified shared experience, termed 'communitas,' as a core component of the ontology of theater (Turner 1969). As such, theater has a long tradition of highly developed techniques for fostering shared experience. In addition, theater has a tradition of bringing the non-human 'to life,' as in puppetry and mask work (Bell 2001). Theater know-how about bringing humans and non-humans into meaningful relation via shared experience may be of use for the development in HMI design strategies, particularly when invited into interdisciplinary research collaboration early in the process, to allow for more meaningful exchange. We plan an exploratory pilot study where we use an interactive theatre approach to develop a set of dramatic scenarios, with the aim of identifying strategies for creating a sense of shared experience in human-machine relations, in the context of a self-driving car simulator. Specifically, using a Wizard of Oz setup we plan to develop scenarios aimed at providing varying degrees of shared experience with an autonomous vehicle. We will then have participants work with the faux autonomous vehicle in order to address a situation with an ethical component. We expect to gain clearer insights into how shared experiences shape trust in HMI and methodological limits/possibilities of the general approach. In the remainder of this abstract we will 1) introduce the recent discussion of reciprocity and shared experience in HMI, 2) discuss the intersection of theater and science, and 3) discuss the concept of morally permissible actions in the context of autonomous vehicle research.

Human-Machine Collaborators as Intersubjective Partners

Andrea Gaggioli et. al. (2021) make a compelling argument for reframing machine collaborators as partners instead of tools: "Effective collaboration in humans stems from 'growing together,' that is, from building a mutual understanding, which evolves over long periods through shared experiences. Consistently, we argue that the introduction of effective robot collaborators hinges upon the development of an intersubjective space between humans and machines" (Gaggioli 2021, 358). Gaggioli et. al. emphasize the machine (in their case, "robotic agent") as already a cultural artifact, acknowledging the creative intersubjective interplay between humans and objects always already in play. While it is acknowledged that humans often take an intentional stance towards non-human entities, the ways in which the intentional stance allows for the reflexive power of such entities to influence humans is often ignored. Following Gaggioli et. al., a dialogic experience structure may be most effective in fostering intersubjective relations between humans and machines (Gaggioli 2021). Theater often features dialogic, narrative, and playful structures, and therefore a theatrical approach may be well-positioned to investigate intersubjective shared experience design for humans and machines.

Theater as Laboratory

Theater has an interesting history as a laboratory for innovation and invention. As a particularly plastic art, theater has the capacity to both incorporate and explore insights from scientific and technology disciplines in active ways, both through incorporating innovations in these fields to production techniques and by engaging core questions from these fields with audiences (Baugh 2014; Rouse 2018). The Bauhaus theater works from the early 1920s were used to explore larger issues about relationships between humans, technology, and affect while also innovating the theatrical form itself (Gropius and Wensinger 1961). Other well-known examples include the landmark "9 Evenings" collaboration between artists and Bell Labs engineers led by Billy Klüver and Robert Rauschenberg in 1966, and more recent work by the Blast Theory collective at University of Nottingham (Benford and Giannachi 2011). These types of art and science collaborations in the theatrical context have a long history which can be traced back at least to early medieval examples (Hoxby 2019, 161). Today, theater is often at the cutting edge of innovation in uses of new technologies, meaning theater today is often a fusion of the live and the digital (Dixon 2007). This positions theater as a medium uniquely suited to reflect, (re)enact, and share our own postdigital lived realities, in which our existence is marked by a heightened awareness of our always shifting entanglement of human, technological, and nonhuman in interconnected negotiation (Haraway 2016).

Autonomous Vehicles and Moral Permissibility

Discussions of human-machine relations often include questions of ethics, such as, how transparent should machine goals be to human interactors (Zonca & Sciutti 2021), and what is the role of encoded bias (Gitelman 2013, Eubanks 2018, Noble 2018, Hamraie 2017). Of particular interest in the current context are ethical issues relating to the kinds of actions a machine, the autonomous car in particular, will take in certain situations. Trolley problems typically take center stage in this kind of discussion, but in many of these types of scenarios the vehicle will not have time to interact with the person, and, on many accounts, the decision will be predetermined algorithmically. For this reason we aim to explore cases of morally permissible action, which are much more likely to occur and provide opportunity for reciprocal HMI. As in William Stafford's poem above, a driver may be faced with several ethical dilemmas involving moral permissibility. Initially the driver in the poem is faced with at least 2 morally permissible options: 1) stop and move the animal, or 2) continue on. While many might prefer that the driver stop and move the animal, it is unclear if such action would be morally obligated. In many cases, it is at least permissible to pass the animal and carry on. Later in the poem, the driver is faced with the decision to save an animal or allow it to die, again a case of morally permissible options. These kinds of scenarios are far more likely to be opportunities for interaction between future drivers and autonomous vehicles. An even more common situation is choosing between different routes which may have a slightly bigger ecological impact, may result in shorter drive times, may affect the chances of congestion, and so on. In this kind of situation, we do not necessarily want to succumb to algorithmic morality and yet the decision has an inescapable ethical dimension (Parvin 2017, 317). A possible interactive solution could conceive of the vehicle as a kind of interactive expert system, offering suggestions and working through the reasons for and against a particular option. In this scenario, there will be a need for mutual trust, perhaps built on shared experiences traveling together. As an initial exploration of this traveling together, we envision an interactive performance which provides a space of shared imagining. In this case theatre can provide a platform for technologists, artists, and the public to come together in order to imagine possible futures for human-machine relations, and provide insights into an experience with possible intelligent HMI partners.

Conclusion: Experimental Performance Design

Using an immersive car simulator platform and Wizard of OZ (WoZ) methods, we plan to run a prototype interactive performance experience as a pilot study in identifying design strategies for shared experience in human-machine relations. We will construct a set of scenarios in which participants will have an opportunity to ride in an autonomous vehicle simulator and interact with a candidate AI, played by an actor. Participants may be with a more or less interactive agent with the aim of creating varying levels of shared experience. Towards the end of the ride the participant will face a scenario in which ethically valanced actions must be taken but which may involve several morally permissible choices. Exit interviews will allow for a clearer picture of participant experiences and guide development of a more rigorous scientific exploration. In line with research ethics guidelines, consent procedures will be followed and a thorough debriefing, including a clear discussion of the WoZ method employed will be provided (Rosén et al., 2021). In addition to providing initial research into design strategies for HMI shared experience, our project will also contribute knowledge toward methods for interdisciplinary art and science research collaborations.

References

Baugh, C. (2014). *Theatre, Performance and Technology: the development and transformation of scenography*. Macmillan.

Benford, S., Giannachi, G. (2011) *Performing Mixed Reality*. MIT Press.

Dixon, S. (2007). *Digital Performance: a History of New Media in Theater, Dance, Performance Art, and Installation.* MIT Press.

Eubanks, V. (2018) *Automating Inequality: How high-tech tools profile, police, and punish the poor.* St. Martin's Press.

Gaggioli, A., Chirico, A., Di Lernia, D., Maggioni, M. A., Malighetti, C., Manzi, F., Marchetti, A., Massaro, D., Rea, F., Rossignoli, D., Sandini, G., Villani, D., Wiederhold, B. K., Riva, G., & Sciutti, A. (2021). Machines Like Us and People Like You: Toward Human–Robot Shared Experience. *Cyberpsychology, Behavior, and Social Networking*, *24*(5), 357–361. https://doi.org/10.1089/cyber.2021.29216.aga

Gitelman, L. Ed. (2013) *Raw Data is an Oxymoron.* MIT Press.

Hamraie, A. (2017). *Building Access: universal design and the politics of disability.* University of Minnesota Press.

Haraway, D. (2016) *Staying with the Trouble: Making Kin in the Chthulucene.* Duke University Press.

Hoxby, B. (2019). Technologies of Performance: From Mystery Plays to the Italian Order. In: *A Cultural History of Theatre in the Early Modern Age*. Bloomsbury. 161-182.

Mou, W., Ruocco, M., Zanatto, D., & Cangelosi, A. (2020). When Would You Trust a Robot? A Study on Trust and Theory of Mind in Human-Robot Interactions. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 956–962. https://doi.org/10.1109/RO-MAN47096.2020.9223551

Noble, S. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Parvin, N. (2017) Our Bodies in the Trolley's Path, or Why Self-Driving Cars Must Not be Programmed to Kill. *Science, Technology & Human Values* 43(2) 302-323.

Rosén, J., Lindblom, J., Billing, E., & Lamb, M. (2021). *Ethical Challenges in the Human-Robot Interaction Field*. The Road to a successful HRI: AI, Trust and ethics - TRAITS Workshop, in conjunction with the 2021 ACM/IEEE International Conference on Human-Robot Interaction, Boulder, USA, March 8--12 2021.

Rouse, R. (2018) Partners: Human and Nonhuman Performers and Interactive Narrative in Postdigital Theater. *Interactive Digital Storytelling: LCNS 11318*. Springer. 369-382.

Sciutti, A., Mara, M., Tagliasco, V., & Sandini, G. (2018). Humanizing human-robot interaction: On the importance of mutual understanding. *IEEE Technology and Society Magazine*, *37*(1), 22–29. https://doi.org/10.1109/MTS.2018.2795095

Turner, Victor. 1969. The Ritual Process: Structure and Anti-Structure. Chicago: Aldine Press.

The Overselling of Super-intelligence: Or, Why Skynet (Probably!) Isn't Taking Over Any Time Soon

Joel Parthemore

Department of Cognitive Neuroscience and Philosophy, University of Skövde

joel.parthemore@his.se

Much concern has been raised in recent years by respectable voices about the potentials and perils of machine superintelligence – what David Chalmers (2010) calls "AI++", which he sets in the context of the so-called *technological singularity*. The Future of Humanity Institute's Nick Bostrum (2014) believes that superintelligence poses one of the primary existential risks to the human species. The matter is urgent: in 1998, he predicted that machine superintelligence would arrive within the first third of the 21st Century – a claim he repeated in (Bostrum, 2005). The late Stephen Hawking suggested that super-intelligent AI could be one of the worst mistakes that the human species will ever make (Kharpal, 2017). Researchers in machine ethics (e.g., Wallach & Allen, 2009) worry about instilling robots with the "right" moral values. The concern is often phrased like this: super-intelligent machines of our creation might look at us the same way we look at ants or even, say, microorganisms; would human morals even make sense to such machines? What if they decided to wipe us out the way the human species decided to eradicate smallpox?

Setting aside the question of whether the existential risk posed by super-intelligent machines really warrants more attention than the arguably more immediate existential risk posed by the triangulation of climate change, mass species extinction and general environmental degradation – all of which scientists are fairly certain are happening now as opposed to in some hypothetical future – the concerns being raised rely on a number of generally unstated yet highly debatable assumptions.

The first is that we know what intelligence is in the first place: something that, so far as I can tell, most researchers in this area take for granted. It is further assumed that we know how to quantify it. One need not buy into Howard Gardner's version of multiple intelligences (Gardner, 1993; Gardner & Hatch, 1989) to allow that IQ tests – the most widely used tests for general intelligence – have known cultural biases and that, even if these biases could be reliably eliminated, one risks reducing a complex, multidimensional phenomenon to a simple linear scale: a concern raised by Steve Torrance (2012, p. 492).

Consider geniuses: geniuses are, after all, not geniuses at everything, and strong anecdotal evidence suggests that those who are labelled geniuses consistently have deficits in certain areas of their cognitive lives. If defining and quantifying intelligence is already a problem, then that of doing so with super-intelligence is even more so. The working definition appears to be "whatever the best of our geniuses have now, only much more so", and vague allusions are made to ants or other presumed-to-be-simple organisms (implying that super-intelligent machines will not just have a whole lot more intelligence but perhaps be intelligent on an entirely different order: that is, the difference will be not just quantitative but qualitative).[1]

The second assumption is that at least in principle we know how to create "true" human or human-level intelligence *now* – other than by the time-tested method of producing children. Indeed, the assumption is often that we have produced at least limited artificial *human* intelligence already. This paper argues that that latter assumption relies on a confusion between two longstanding traditions in AI research – one of which is interested in engineering practical solutions to the exclusion of philosophical considerations about intelligence, the other of which is interested in creating the functional equivalent of living minds – and on the faulty reasoning that, if computers are so much better than human beings at performing many of the seemingly intelligent tasks they perform now, there is nothing to stop them exceeding human capacities across the board.

There seems little denying that the AI-as-engineering community has racked up astounding successes – leading researchers like Blay Whitby (2003) to complain that the goalposts of what count as intelligence keep getting shifted so that they are always "what machines have not been able to do yet". The argument can and has been made though that all these successes prove is that a great many tasks that were assumed to require sophisticated cognitive capacities can, in fact, be reduced to relatively simple algorithms. Indeed, some have taken this line of

[1] Indeed, this is how David Chalmers (2010) distinguishes AI+ (slightly smarter than human) from AI++.

thought to the conclusion that *all* of human intelligence can mechanized this way – reduced to mathematical descriptions – even that human beings are a kind of self-deluding automata; even as others – Roger Penrose (1994: pp. 72-77) comes particularly to mind – argue that key aspects of human intelligence are not mathematically describable at all, exceeding the capacities of *any* formal system. This is to say that the human mind almost certainly *can* be described by a formal system[2] but one that is an order of magnitude more expressively powerful than that used in current computer systems.

This paper uses the discussions over super-intelligence to argue that the best answer *for now* for what human intelligence is lies somewhere between these two extremes. On the one hand, key aspects of human intelligence are not describable in the mathematics that describe essentially all current computer systems; on the other, this does not mean that these aspects of human intelligence cannot be captured mathematically.

In the formal systems that drive current computer systems,[3] any version of

$$p \ \wedge \neg p$$

…renders the entire system inconsistent. Putting this another way, there is a fairly simple set of self-referential statements (the mathematical equivalent of Epimenides' paradox) that cannot be accommodated in a reliable and consistent manner because they appear to challenge the binary distinction of truth-functional statements into *either* true *or* false.[4] That renders attempts to address inconsistent reasoning in contemporary computer systems limited and essentially *ad hoc.*

Human beings, on the other hand, seem capable of tolerating (even exploiting) inconsistencies systematically without becoming globally inconsistent (meaning that they can hold contradictory beliefs in multiple areas while otherwise continuing to engage in rational thought and reason accurately). This allows them to entertain propositions that simply could not be entertained before – and also to make what a rational observer might or even reasonably would take to be a mistake. How?... by taking context into account. One can entertain the proposition p in one context (A) and the proposition $\neg p$ in another (B). In effect, propositions can be annotated by context:

$$p^{\mathrm{A}} \ \wedge \neg p^{\mathrm{B}}$$

…Which no longer presents an obvious contradiction except in the case where $A = B$. Such a context-sensitive system is, naturally, phenomenally more complex and so difficult if not indeed impossible to comprehend in its entirety. On the other hand, it can be built; and human beings have good experience of creating things that ultimately outstrip human capacity to explain them.

If this line of reasoning is at all correct, then human intelligence is (in certain critical respects at least) far more complex than it has generally been given credit; while the road to super-intelligence – if the term "super-intelligence" can be seen potentially to make sense *in the future* – lies through arriving at a better understanding of human (and, doubtless, non-human animal) intelligence first. If, at some point, we succeed in building artificial human-level intelligences (sometimes called *artificial general intelligences*), we may find that they are subject to many of the same cognitive limitations we are, making many of the to-date distinctly human mistakes that human beings make.

References

Bostrum, N. (2014). *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press.

Bostrum, N. (2005). How long before superintelligence? *Linguistic and Philosophical Investigations, 5*(1).

Chalmers, D.J. (2010). The singularity: A philosophical analysis. *Journal of Consciousness Studies, 17*(9-10): 7-65.

[2] Saying alternatively that it is "driven by a formal system" would introduce ontological concerns that are probably best avoided.

[3] …My point being that there are different species of formal systems with different levels of expressive power.

[4] Of course, as Kurt Gödel (1992) is widely accepted as having proven, effectively *all* formal systems of sufficient expressive power to express any kind of self-referential statement are unable to express *some* class of self-referential statements without becoming inconsistent in the process: i.e., formal systems, by their nature, cannot be both complete and consistent. There will be statements that can be expressed in the formal system whose truth value cannot be determined within that system.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. Basic Books.

Gardner, H. & Hatch, T. (1989). Educational implications of the theory of multiple intelligences. *Educational Researcher*, *18*(8): 4-10. https://doi.org/10.3102%2F0013189X018008004

Gödel, K. (1992). *On formally undecidable propositions of* Principia Mathematica *and related systems*. Courier (Dover Books).

Kharpal, A. (2017). Stephen Hawking says A.I. could be "worst event in the history of our civilization". CNBC, 6 November 2017. Accessed 27 October 2021 at https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html

Penrose, R. (1994). *Shadows of the mind: A search for the missing science of consciousness.* Oxford University Press.

Torrance, S. (2012). Super-intelligence and (super-)consciousness. *International Journal of Machine Consciousness*, *4*(2): 483-501.

Wallach, W. & Allen, C. (2009) *Moral machines: Teaching robots right from wrong.* Oxford University Press.

Whitby, B. (2003) The myth of AI failure. *Cognitive Science Research Papers 568.* University of Sussex. Accessed 27 October 2021 at http://www.sussex.ac.uk/informatics/cogslib/reports/csrp/csrp568.pdf

Ecosystems of reinforcement learning agents

Claes Strannegård*

Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
Sweden

Ecosystem models play an important role in analyzing the ecological consequences of human activities such as fishing, hunting, farming, logging, and construction. What are the ecological consequences of constructing a highway through a certain forest or allowing fishing in a certain marine area? Ecosystem models can be used for answering such questions and thus contribute actively to sustainable decision-making. Interestingly, many of the most used ecosystem models do not model animal cognition at the individual level at all, or leave out some fundamental mechanisms of adaptation. For example, the widely used model Ecopath [1] leaves out learning, while Google DeepMind's model [2] leaves out evolution. This extended abstract briefly presents the agent-based open-source ecosystem model Ecotwin [3], which was constructed with the aim of bridging this apparent gap. The features of Ecotwin include individual lifelong reinforcement learning, reward signals based on homeostasis, reproduction based on physical encounters, and terrain models based on geographic 3D data. The agents are equipped with nervous systems consisting of four integrated neural networks that compute, respectively, reflexes, rewards, next actions, and next states.

1 Ecosystem models

Among the oldest examples of ecosystem models are the Lotka-Volterra differential equations for predator-prey population dynamics [4], in which animals are modeled at an aggregated level, e.g. as the number of individuals of a certain species. Thus, there is no model of individual animals and in particular no model of animal cognition involved. A related approach uses similar equations together with a terrain map that is divided into cells with their own local populations. An example is the Ecopath/Ecosim/Ecospace simulator for marine ecosystems [1]. Another approach models animals at the individual level and studies the development of reflex agents according to evolutionary algorithms [5]. Still another approach uses populations of pure reinforcement learning (RL) agents [2, 6]. These agents typically respawn at random locations after death, with the knowledge learned during previous lives passed on to the next generation. Several ecosystem models are seriously unrealistic, since they omit fundamental mechanisms of animal reproduction and cognition. This limits their applicability for predicting the development of real ecosystems, with or without human

* Corresponding author: claes.strannegard@chalmers.se

interventions. More realistic ecosystem models might include features such as these:

– Genotypes and phenotypes
– Sexual and asexual reproduction
– Sexual reproduction based on physical encounters
– Models of animals and plants
– Death criteria based on homeostasis
– Modularized model of nervous systems
– Reward based on multiple homeostatic signals
– Positive and negative reflexes
– Individual and lifelong RL
– 3D terrain model

2 Ecotwin

As far as I know, no ecosystem model satisfies these features. Ecotwin was designed with the aim of filling this gap. Ecotwin models animal cognition via a rough model of a nervous system. Nervous systems are prevalent in almost all taxa of the animal kingdom. In contrast to standard deep reinforcement learning models, nervous systems are typically not monolithic, but highly modular. Static circuits such as reflexes may appear side-by-side with plastic circuits that develop with RL [7, 8]. Of particular interest here are the following structures of the nervous system, which are shared by many animal species:

– the *brain stem and reflex circuits*, controlling reflexes —positive, e.g., the suck reflexes and the knee reflex, and negative, e.g., the diving reflex [7]. Reflexes can save time and even lives. In particular, they can prevent animals from dangerous exploration, e.g., jumping from high cliffs or inhaling water.
– the *basal ganglia*, which maps sensory signals to actions [9]
– the *insula*, which maps internal signals, such as glucose level and external signals, such as pheromone smell, to signals linked to reward [10]
– the *medial prefrontal cortex*, which predicts the outcomes of actions [11].

These structures are modeled as follows in Ecotwin:

Definition 1. *A nervous system consists of:*

– *a set of sensors* S
– *a set of actions* A
– *a reflex network with input nodes* S *and output nodes* A
– *a policy network with input nodes* S *and output nodes* A
– *a happiness network with input nodes* S *and a single output node for representing a scalar happiness value, and*
– *a prediction network with input nodes* $S \cup A$ *and output nodes* S

The reflex and happiness networks remain the same throughout the individuals' lives. In contrast, the policy and prediction networks are updated continuously. The policy network is updated via RL. The reward signal at time t is defined as $happiness(t) - happiness(t-1)$, where *happiness* is computed by the happiness network. For model-based algorithms, the prediction network can be used as a model. The prediction network is updated continuously by means of supervised learning. The above definition covers pure reflex agents (with an empty policy network), pure RL-agents (with an empty reflex network), and agents that combine reflexes with RL. By leaving the nervous system empty, one may also model organisms lacking nervous systems altogether, such as bacteria and plants.

3 Results

Ecosystem models should be evaluated not by their structure, but by their ability to reproduce animal behavior and ecosystem development as observed in real ecosystems. Ecotwin was used to model a three-species terrestrial ecosystem in which Lotka-Volterra style predator-prey dynamics emerged and a marine ecosystem in which a diel vertical migration pattern emerged [3]. Ecotwin is open source and available for download at `www.ecotwin.se`.

References

1. V. Christensen and C. J. Walters, "Ecopath with Ecosim: methods, capabilities and limitations," *Ecological modelling*, vol. 172, no. 2-4, pp. 109–139, 2004.
2. P. Sunehag, G. Lever, S. Liu, J. Merel, N. Heess, J. Z. Leibo, E. Hughes, T. Eccles, and T. Graepel, "Reinforcement learning agents acquire flocking and symbiotic behaviour in simulated ecosystems," in *Artificial life conference proceedings*, pp. 103–110, MIT Press, 2019.
3. C. Strannegård, N. Engsner, P. Ferrari, H. Glimmerfors, M. H. Södergren, T. Karlsson, B. Kleve, and V. Skoglund, "The ecosystem path to general AI," *arXiv preprint arXiv:2108.07578*, 2021.
4. A. J. Lotka, *Elements of Physical Biology, by Alfred J. Lotka*. Williams & Wilkins, 1925.
5. C. G. Langton, "Artificial life: An overview," 1997.
6. J. Yamada, J. Shawe-Taylor, and Z. Fountas, "Evolution of a complex predator-prey ecosystem on large-scale multi-agent deep reinforcement learning," in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2020.
7. R. B. Stein and C. Capaday, "The modulation of human reflexes during functional motor tasks," *Trends in neurosciences*, vol. 11, no. 7, pp. 328–332, 1988.
8. E. O. Neftci and B. B. Averbeck, "Reinforcement learning in artificial and biological systems," *Nature Machine Intelligence*, vol. 1, no. 3, pp. 133–143, 2019.
9. L. Ding and J. I. Gold, "The basal ganglia's contributions to perceptual decision making," *Neuron*, vol. 79, no. 4, pp. 640–649, 2013.
10. R. Nieuwenhuys, "The insular cortex: A review," in *Evolution of the Primate Brain* (M. A. Hofman and D. Falk, eds.), vol. 195 of *Progress in Brain Research*, pp. 123–163, Elsevier, 2012.
11. W. H. Alexander and J. W. Brown, "A general role for medial prefrontal cortex in event prediction," *Frontiers in computational neuroscience*, vol. 8, p. 69, 2014.

Remote and lab-based research of viewing experiences and recollection in cinematic virtual reality

Kata Szita[1], Pierre Gander[2]

[1]*Department of Applied IT, University of Gothenburg; Trinity Long Room Hub, Trinity College Dublin*

[2]*Department of Applied IT, University of Gothenburg*

Presenter's e-mail address: kata.szita@gu.se

Introduction

Cinematic virtual reality (VR) affords 360-degree viewing experiences during which a viewer's body position defines the momentary viewing perspective: a moving-image content is presented in 360 degrees, while one only accesses the segment of the narrative space toward which they are facing at any given time (Mateer, 2017). The 360-degree space prompts highly immersive experiences and embodied engagement (Van den Broeck, Kawsar, & Schöning, 2017). However, the arbitrariness and constant changes of viewing perspectives may affect a viewer's comprehension and recollection of narrative elements. The study proposes that such immersive experiences and sensorimotor involvement induce a first-person perspective to observe narrative events (St. Jacques, 2019). As opposed to a camera perspective, first-person perspective is associated with an increased sensation of narrative presence and emotional engagement, as well as more accurate and vivid recollection of information (Serino & Repetto, 2018). To determine these effects, we designed an experiment to compare viewers' reactions to an animated movie, *Pearl* (Osborne, 2016), either watched using a head-mounted display or on a regular screen.

Due to the restrictions related to the Covid-19 pandemic, laboratory testing of the hypothesis was no longer feasible, therefore an online experiment was designed. In this experiment, participants were recruited from internet-based communities and were requested to use their own equipment at a location of their choice.

In the following, we present the methods and findings of the online experiment (Szita, Gander, & Wallstén, 2021) and the methodological challenges and limitations remote data collection holds. Based on these limitations, we outline a follow-up experiment to be conducted in laboratory settings.

Method

To isolate the effects of cinematic virtual reality, we followed a between-subjects design comparing virtual reality and screen viewing of the same six-minute film. *Pearl* has a 360-degree and a screen-based version. In the VR format, the viewer can define the direction of viewing within the six degrees of freedom that allows for a full range of head movements. The screen version uses multiple camera angles.

The study involved 165 participants aged 16–62 (M = 30.44, SD = 9.61). The criterion for taking part was access to the respective screening appliance (tethered VR headset or a stationary screen of minimum 12 inches). Eighty-five participated in the VR condition and 80 in the screen condition. The VR version was advertised in online groups for virtual reality users to target users who have access to headsets and are experienced with VR technology. This was to avoid novelty experiences' biasing effects and to assume that each participant uses the most suitable settings (e.g., interpupillary distance, screen resolution). Participants for the screen version were recruited through online channels targeting communities of cinema enthusiasts and general audiences.

After watching the respective movie sequence, participants were asked to complete an online survey using Psytoolkit (Stoet, 2017). The survey consisted of three sections measuring viewing experience, memory characteristics, and recollection accuracy, and an additional set of questions recorded demographic data, user habits, and technical details of participation (e.g., VR headset type).

The first section (ten items) measured emotional engagement, sense of presence in the fictional space, empathy toward characters, awareness of the physical surroundings, and physiological reactions (e.g., nausea, dizziness) (based on Fonseca & Kraus, 2016; Zhang, 2020). The second section, memory characteristics (ten items), recorded recollection vividness, emotional and physical reactions when recalling the movie, memory perspective (first- or third-person), and the structural comprehension of the narrative (Berntsen & Rubin, 2006; Qin, Rau, & Salvendy, 2009). The third section measured recollection accuracy (Syrett, Calvi, & van Gisbergen, 2016; Szita & Rooney, 2021): participants were given twelve statements from the movie to determine whether they were true or false.

Results and discussion

The values for each of the dependent variables were not normally distributed, therefore, we used a Mann–Whitney U test to compare the two conditions. The variables that showed significant differences are shown in Table 1.

Table 1. Mann-Whitney U test: mean rank and median values, z-scores, and effect sizes

Variable	Scale (from 1 to 7)	VR		Screen		z-score	r^2
		Mean rank	Median	Mean rank	Median		
Viewing experience							
I felt I was inside the story**	not at all–completely	92.68	5	72.71	5	-2.74	.046
I felt I was at the places in the displayed environment***	not at all–completely	102.93	5	61.83	4	-5.62	.191
When I was watching, time seemed to fly quickly***	not at all–completely	95.96	5	69.23	4	-3.65	.081
This experience was fascinating***	not at all–completely	97.45	6	67.65	4.5	-4.08	.101
I felt nauseous while watching the movie*	not at all–completely	88.74	1	76.91	1	-2.12	.027
Memory characteristics							
When I think of the movie, I can see with my mind's eye what took place*	not at all–as clearly as if I watched it now	89.92	6	75.65	5.5	1.98	.024
The relative spatial arrangement of people and objects in my memory for the movie is***	vague–clear	96.64	6	68.51	5	3.93	.094
At parts, the movie made me feel happy. This feeling was*	weak–strong	90.54	6	74.99	5	2.13	.027
When I now recall the movie, I primarily see what happened from a perspective as seen from**	inside the story world–as an outside observer	72.43	5	84.23	6	-3.00	.055
Recollection accuracy							
Percentage of correct answers**	–	72.06	66.7	94.63	75	-3.07	.057

Note: * $p < .05$, ** $p < .01$, *** $p < .001$

Ratings for feeling like being inside the story and in the displayed environment were significantly higher in the VR condition than the screen condition (U = 2577, U = 1706). Measuring detachment from the physical environment, ratings showed significantly higher values in the case of VR (U = 2298.5). Additionally, participants felt more fascinated by the VR experience than screen viewing experience (U = 2172) but also felt more nauseous in VR than during screen viewing (U = 2912.5). Thus, although VR viewers were found to be more immersed—which may have been facilitated by the fact that participants were recruited from communities gathering active VR users—they were also more likely to experience cybersickness. Discomfort, such as cybersickness, may draw attention to one's physical body, thereby hindering immersion. This might be the reason for the lack of significant differences between screen and VR viewers' emotional engagement and empathy with characters.

VR participants were more likely to recall narrative events as clearly as if they watched the movie at the moment of answering than screen viewers (U = 3988). Correspondingly, the relative spatial arrangement of people and objects in participants' memory was rated clearer in the VR condition (U = 4559). VR viewers reported a stronger feeling of happiness when recalling the movie (U = 4040.5). To measure first- versus third-person recollection perspectives, participants rated their experience on a scale stretching from "inside the story world" to "as an outside observer looking into the story world." Supporting our hypothesis, VR viewers reported recollection more from inside the story world through a first-person perspective than screen viewers (U = 2501.5). In terms of recollection accuracy, screen participants recalled the movie more accurately than VR participants (U = 2470).

We found that VR viewing would more likely induce a first-person point of view while screen viewing leads to a third-person perspective. Although participants rated their memories of the movie clearer in the VR condition, the accuracy of recollection was poorer signaling that the attributes of VR help engagement but hinder access to all visual details. In other words, our results suggest a causal relationship between the 360-degree field of simulation and attention: VR viewers need to turn their bodies to access information in the different parts of the visual field which may cause them to miss details that are momentarily obscured.

Limitations and future work

This study was conducted in natural settings; each participant watched the movie on their own device and in an environment of their choice. Such a natural experiment leads to results with high ecological validity as participants followed their general routines and our detailed instructions allow for replicability. But while online data collection omits geographic constraints, we were unable to control eventual extraneous variables, such as viewing environments and distractions. Additionally, although using one's personal device and settings would likely lead to a comfortable viewing experience, we cannot rule out the bias of individual devices (e.g., differences in field of view or resolution). Therefore, a laboratory study to confirm our results is an informative next step to make conclusions of the causality of viewing conditions. It would also allow for using a participant pool of both experienced and inexperienced users irrespective of access to virtual reality headsets.

References

Berntsen, D., & Rubin, D. C. (2006). Emotion and vantage point in autobiographical memory. *Cognition and Emotion*, *20*(8), 1193–1215. DOI: 10.1080/02699930500371190

Fonseca, D., & Kraus, M. (2016). A comparison of head-mounted and hand-held displays for 360° videos with focus on attitude and behavior change. *Proceedings of the 20th International Academic Mindtrek Conference* (287–296). Mindtrek. DOI: 10.1145/2994310.2994334

Mateer, J. (2017). Directing for cinematic virtual reality: How the traditional film director's craft applies to immersive environments and notions of presence. *Journal of Media Practice*, *18*(1), 14–25. DOI: 10.1080/14682753.2017.1305838

Qin, H., Rau, P.-L. P., & Salvendy, G. (2009). Measuring player immersion in the computer game narrative. *International Journal of Human–Computer Interaction*, *25*(2), 107–133. DOI: 10.1080/10447310802546732

Serino, S., & Repetto, C. (2018). New trends in episodic memory assessment: Immersive 360° ecological videos. *Frontiers in Psychology*, *9*, 1878–1878. DOI: 10.3389/fpsyg.2018.01878

St. Jacques, P. L. (2019). A new perspective on visual perspective in memory. *Current Directions in Psychological Science*, *28*(5), 450–455. DOI: 10.1177/0963721419850158

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*(1), 24–31. DOI: 10.1177/0098628316677643

Syrett, H., Calvi, L., & van Gisbergen, M. (2016). The Oculus Rift film experience: A case study on understanding films in a head mounted display. In R. Poppe, J.-J. Meyer, R. Veltkamp, & M. Dastani (Eds.), *Proceedings of The International Conference on Intelligent Technologies for Interactive Entertainment* (197–208). Springer. DOI: 10.1007/978-3-319-49616-0_19

Szita, K., Gander, P., & Wallstén, D. (2021). The effects of cinematic virtual reality on viewing experience and the recollection of narrative elements. *PRESENCE: Virtual and Augmented Reality*, *27*(4), 410–425. DOI: 10.1162/PRES_a_00338

Szita, K., & Rooney, B. (2021). The effects of smartphone spectatorship on attention, arousal, engagement, and comprehension. *i-Perception*, *12*(1), 1–20. DOI: 10.1177/2041669521993140

Van den Broeck, M., Kawsar, F., & Schöning, J. (2017). It's all around you: Exploring 360° video viewing experiences on mobile devices. *Proceedings of the 25th ACM International Conference on Multimedia* (762–768). Association for Computing Machinery. DOI: 10.1145/3123266.3123347

Zhang, C. (2020). The why, what, and how of immersive experience. *IEEE Access*, *8*, 90878–90888. DOI: 10.1109/ACCESS.2020.2993646

Escaping 'Death by GPS': Foundations for Adaptive Navigation Assistance

Fateme Teimouri[1], Kai-Florian Richter[1]

[1]*Department of Computing Science, Umeå University*

teimouri@cs.umu.se, kaifr@cs.umu.se

Introduction

Navigation services usefully combine globally available positioning infrastructure (global navigation satellite systems; GNSS), mobile computing facilities, and geographic data sources to offer wayfinders an anywhere, anytime experience of knowing where they are and where they need to go. No wonder such services have become highly popular and find widespread use, be it as in-built systems in cars or running on mobile devices. However, their use may also lead to various kinds of mishaps and incidents, and most often leave wayfinders clueless about how they ended up at their destination. These issues are so prevalent that a new term has been coined for them– 'Death by GPS'. Sadly, for some wayfinders this term has to be understood very literally (Lin, Kuehl, Schöning, & Hecht, 2017).

While research has identified both technological and–arguably more importantly–cognitive reasons for why people fail in their navigation and do not learn anything about the environment when using these services (e.g., Gardony, Brunyé, Mahoney, & Taylor, 2013), to date there has been only very little research on how to counter these negative effects (e.g., Brügger, Richter, & Fabrikant, 2019). The little research that exists has largely only looked at changes of human behavior and spatial learning given specific manipulations of device interaction or instruction giving. Without doubt this addresses important questions as such research may offer insights into design principles for cognitively more adequate navigation services. But these manipulations are usually manually created and specific to the study environment. This kind of research disregards navigation services' adaptive potential and that, ultimately, they need to provide such assistance automatically and in any environment. In other words, in the interaction triangle of wayfinder, environment and service, the service has been largely ignored so far.

This VR project, funded by the Swedish Research Council (Vetenskapsrådet), aims at identifying where and when navigation services need to adapt their information provision, such that navigation is at least as efficient as with today's services, but wayfinders' understanding of the navigation process increases, and potential disagreement of system and wayfinder on how to best describe a wayfinding situation decreases. In previous work (Teimouri & Richter, 2020), we studied the effects of most current path search models–be it commercial or from research–in some way or another assuming that we are 'alone' while navigating, which has consequences for travel time (we may get stuck in traffic) or social costs (we may suddenly move through somebody's 'backyard'). Further, to increase a wayfinder's understanding of the navigation process, in particular, the route traveled through an environment, we developed an approach for identifying those locations along a route that define its characteristics, which we term *route-defining locations*. In an agent-based simulation we showed that this mechanism is also effective in compressing route instructions to the essential information, which varies depending on a wayfinder's previous knowledge (Teimouri & Richter, 2021).

However, this previous work has not considered yet how well instructions match with human conceptualization of a given wayfinding situation, which is one of the main objectives of this VR project. Thus, here we present an empirical evaluation on how much people agree that some navigation instructions usefully describe a given wayfinding situation. To this end, we collected both quantitative (ratings) and qualitative data (comments and alternative instructions).

Methods

This research is motivated by some anecdotal evidence that there are situations when a navigation service generates instructions that do not fit a wayfinder's understanding of the same situation, despite the instructions being technically correct (cf. Hirtle, Richter, Srinivas, & Firth, 2010). An online questionnaire was created to evaluate such potential mismatches. The questionnaire was shared using social media (thus, the actual identity of participants is unknown). It is divided into two sections. First, it gathers some basic demographic information that can be used to perform aggregate data analysis. Then, it shows a series of 18 screenshots from Google Street View of various navigation situations, along with the matching instructions generated by Google Maps for each situation. These screenshots depict various driving situations in several countries that were selected by the authors; each was presented on its own screen (see *Figure 1*). In total, 46 anonymous participants took part in the study who were mostly men (27 males, 17 females, and 2 others), with the majority holding a driver's license (39 with, 7 without). The participants' task was to express how much they agree that the given instructions accurately reflect the situation shown. The scale ranges from 1 to 5, with 1 indicating ‘strongly disagree’ and 5 indicating ‘strongly agree.’ With each situation, participants also had the option of providing an alternate instruction that, in their opinion, better fits the scenario by entering text in a text field.

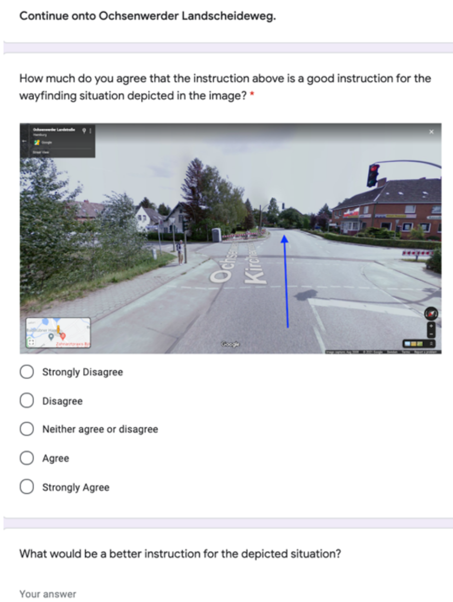


Figure 1 An example screen taken from the questionnaire.

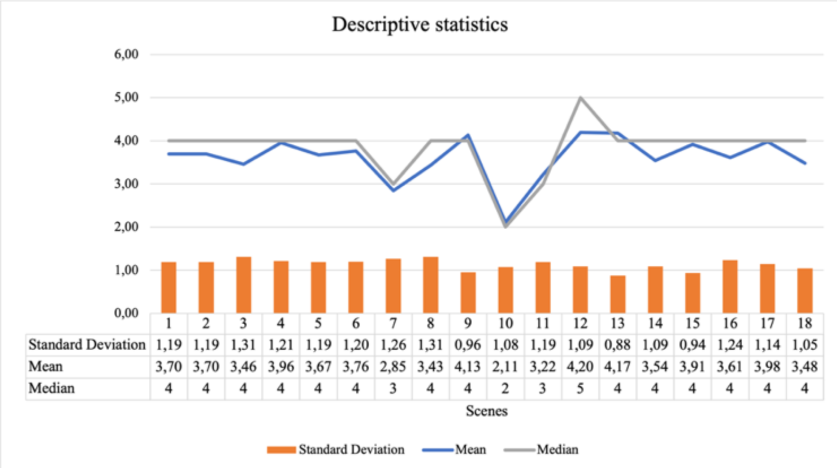
Results

Figure 2 shows mean and median ratings for each scene. As we can observe, scene 7 and scene 10 received the lowest ratings, which indicates that sometimes instructions seem to be not quite right, i.e., not match with human understanding. Such mismatches can lead to ambiguity about what to do, increase cognitive load and decision-making time, lower a wayfinder’s trust in the system, and, in general, cause navigational errors and risky actions (cf. Lin, Kuehl, Schöning, & Hecht, 2017).

Qualitative data, i.e., the collected comments and alternative instructions from the participants, provide a corpus of human-generated instructions. These modified instructions also allow for determining what problems participants had with the original instructions, i.e., to discover patterns (or sources) of incompatibilities between the service's instructions and human understanding of a scenario. These comments can be grouped into two categories: 1) issues arising from the language used to verbalize instructions, and 2) issues arising from the representation utilized to infer the navigation actions to perform.

Discussion

Our results seem to clearly indicate that indeed some issues exist regarding the instructions provided by current navigation services. We identified the language used and the underlying representation as two main causes for these issues. To overcome the language issues, navigation services should adjust and enhance their verbalization capabilities to better align with user expectations and preferences. Issues related to mismatches between how a navigation service describes a scene and how wayfinders may interpret it appear to be the more challenging task to solve. Karimi, Roongpiboonsopit, & Kasemsuppakorn (2011) discussed five sources of uncertainty in navigation services, which stem from: 1) the road network base data; 2) converting addresses into geographic coordinates; 3) estimating geographic coordinates of a person or vehicle's current location; 4) determining the current location of the vehicle on the road network; and 5) computing step-by-step instructions on how to travel on routes using an inference system. The most challenging of these issues, and the one we believe mostly at play in our study, is the final one.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Standard Deviation	1,19	1,19	1,31	1,21	1,19	1,20	1,26	1,31	0,96	1,08	1,19	1,09	0,88	1,09	0,94	1,24	1,14	1,05
Mean	3,70	3,70	3,46	3,96	3,67	3,76	2,85	3,43	4,13	2,11	3,22	4,20	4,17	3,54	3,91	3,61	3,98	3,48
Median	4	4	4	4	4	4	3	4	4	2	3	5	4	4	4	4	4	4

Figure 2: Descriptive statistics for all 18 scenes.

While the current study has certain limitations–among them the static nature of the study material–we believe it clearly points to the fact that there are issues with how navigation services provide instructions. Thus, it is just the beginning of a broader investigation into the challenges that arise when a navigation service's inference mechanisms and human conceptualization of wayfinding situations differ. Future work will include more empirical work utilizing a more dynamic presentation of the wayfinding situations (e.g., videos or moving through a VR environment), and implementing a reasoning mechanism that determines how likely it is that a given situation is not adequately captured by the inference system, for example, because some data is not available or because a turn is on the border between two different turn direction concepts. The empirical findings will inform such reasoning mechanisms, which will subsequently be incorporated into instruction production, i.e., instructions will communicate the system's confidence in capturing a situation according to the wayfinder's understanding.

References

Brügger, A., Richter, K.-F., & Fabrikant, S. I. (2019). *How does navigation system behavior influence human behavior?* Cognitive Research: Principles and Implications, 1-22.

Gardony, A., Brunyé, T., Mahoney, C., & Taylor, H. (2013). *How navigational aids impair spatial memory: Evidence for divided attention*. Spatial Cognition & Computation, 319-350.

Hirtle, S., Richter, K.-F., Srinivas, S., & Firth, R. (2010). *This is the tricky part: When directions become difficult*. Journal of Spatial Information Science, 53-73.

Karimi, H., Roongpiboonsopit, D., & Kasemsuppakorn, P. (2011). *Uncertainty in personal navigation services*. The Journal of Navigation, 341-356.

Lin, A., Kuehl, K., Schöning, J., & Hecht, B. (2017). *Understanding" Death by GPS" A Systematic Study of Catastrophic Incidents Associated with Personal Navigation Technologies*. (ss. 1154-1166). Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.

Teimouri, F., & Richter, K.-F. (2020). *You Are Not Alone: Path Search Models, Traffic, and Social Costs.* 11th International Conference on Geographic Information Science. Poznan, Poland.

Teimouri, F., & Richter, K.-F. (2021). *Abstracting Routes to their Route-Defining Locations.* Computers, Environment and Urban Systems, Under Revision.