15TH BIANNUAL CONFERENCE OF THE GERMAN SOCIETY FOR COGNITIVE SCIENCE

5.-7. SEPTEMBER 2022

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

GESELLSCHAFT FÜR KOGNITIONSWISSENSCHAFT

Cognitive Science

Organizing Committee

Evelyn Ferstl (chair), Lars Konieczny, Rul von Stülpnagel, Judith Beck, Lisa Zacharski

Support

Martin Oettle, Marion Winski-Steiert

Program Committee

Judith Beck, Tobias Bormann, Evelyn Ferstl, Adriana Hanulíková, Lars Konieczny, Nele Rußwinkel, Ute Schmid, Frieder Vogelmann, Rul von Stülpnagel, Lisa Zacharski

Student assistents

Leah Bohr, Mateo Cortés, Deborah Ewert, Adrien François, Andrew Heinz, Laura Jenne, Sarah Kapp, Katharina Klotz, Henrik Lorenzen, Rosalie Oetinger, Lars Rudnick, Samuel Sester, Tim Sudermann

Reviewers

Asya Achimova, Nicole Altvater-Mackensen, Benjamin Angerer, Thomas Barkowsky, Clemens Georg Bartnik, Karin Bausenhart, Thea Behrens, Anna Belardinelli, Alexander Bergs, Tarek Richard Besold, Bettina Bläsing, Alice Blumenthal-Dramé, Tobias Bormann, Simon Büchner, Martin Butz, Francesca Capuano, Claus-Christian Carbon, Amory H. Danek, Hannah Deininger, Francesca Delogu, Dominik Endres, Bettina Finzel, Michael Franke, Claudia Friedrich, Klaus Gramann, Fritz Günther, Kai Hamburger, Adriana Hanulikova, Linda Heimisch, Elisabeth Hein, Annette Hohenberger, Markus Huff, Frank Jäkel, Jana B Jarecki, Barbara Kaup, Martin Christof Kindsmüller, Markus Knauff, Stefan Kopp, Jakub Krukar, Pablo Leon Villagra, Katharina Lüth, Kyla McConnell, Franz-Benjamin Mocnik, Julia Müller, Albert Newen, Eva Nuhn, Dario Paape, Ulrike Pompe-Alama, Kathrin Reichmann, Kai-Florian Richter, Florian Röser, Agnes Rosner, Constantin Rothkopf, Nele Rußwinkel, Ute Schmid, Gregor Schöner, Juliane Schwab, Mikhail Spektor, Frieder Stolzenburg, Manfred Thüring, Sabine Timpf, Frieder Vogelmann, Bettina von Helversen, Thomas Wallis, Gert Westermann, Eva Wiese

TABLE OF CONTENTS

Dear participants of KogWis2022,

we warmly welcome you to the 15th Biannual Meeting of the German Cognitive Science Society (Gesellschaft für Kognitionswissenschaft, e. V.). The conference usually takes place every two years but had to be cancelled in 2020 due to the Covid pandemic. It is our great pleasure to finally welcome you back in Freiburg. Many of us are eager to get together again in person and to engage in fruitful and lively discussions.

The Albert-Ludwigs-University in Freiburg appointed in 1991 the first chair of Cognitive Science in Germany, Gerhard Strube, who was instrumental in co-founding the GK in 1994. In the same year, Freiburg hosted the first KogWis. Cognitive Science was one of three areas at the Institute for Computer Science and Social Research (Institut für Informatik und Gesellschaft, IIG). The institute was ahead of its time in studying the impact and the consequences of Computer Science and the emergence of digital technologies on society.

Due to the recent technological advances, issues represented at the IIG, such as ethics, security or diversity, are now more important than ever and have since then permeated the public sphere. Cognitive Science in Freiburg is now widely connected within the university, reflecting our research interests in language and communication, neuroscience, and gender studies.

Cognitive Science ties together seemingly divergent disciplines, in particular psychology, linguistics, computer science, philosophy, anthropology, and neuroscience. With its interdisciplinary focus, the field has the unique potential to contribute to tackling the challenges of digitalization.

The motto of KogWis2022 *Understanding Minds* represents two perspectives: The conference provides a forum for all topics in the study of how minds – both human and artificial – operate. The theme also puts a specific spotlight on how cognitive systems make sense of the world, in particular in language comprehension and communication.

Both the plenary sessions as well as the over 100 individual submissions make up an interesting and stimulating program. We very much hope that you will enjoy the conference and wish you a pleasant stay in Freiburg.

September 2022

Evelyn C. Ferstl
Lars Konieczny
Rul von Stülpnagel
Judith Beck
Lisa Zacharski

Center for Cognitive Science, Department of Psychology, University of Freiburg

Program Overview

Monday

8:00	**Registration**	
9:00	Welcome Address **Aula**	
9:20	Keynote 1 **Marcel Brass** Free Will: An Empirical Perspective **Aula**	
10:20	**Coffee Break**	
	Session 1 **Decision Making & Categorization** Hörsaal 1098	Symposium 1 **Computational Modelling of the Active Self** Hörsaal 1199
10:40	In search of lost memories: modeling forgetful generalization **Breit, Sakkaki, Murayama, Wu**	Computational cognitive modeling of the predictive active self in situated action **Rußwinkel, Kopp**
11:00	Eliciting everyday beliefs using random generation tasks **León-Villagrá, Castillo, Chater, Sanborn**	Tool-use and agency in artificial agent **Hafner**
11:20	Non-symbolical Bayesian reasoning with proportions and probabilities **Loibl, Leunders**	A neural dynamic account of intentionality as the basis of an active self **Schöner, Tekülve**
11:40	Do verbally processed stimulus-affect contingencies establish stimulus-affect associations? **Martiny-Huenger**	Homeostasis drives the active self while generative models constitute it **Butz**
12:00	Discussion	Discussion
12:20	**Lunch Break**	

1:40	KEYNOTE 2 **Matthew W. Crocker** The Neurocomputation of Sentence Meaning AULA	
2:40	**Coffee Break**	
	SESSION 2 **LANGUAGE PROCESSING** HÖRSAAL 1098	SYMPOSIUM 2 **FLEXIBLE REPRESENTATIONS FOR HUMAN(-LIKE) PROBLEM SOLVING** HÖRSAAL 1199
3:00	Framing effects and numeral modification **Claus**	Strategies and representations for solving sudokus **Behrens, Jäkel**
3:20	How morphological family size affects word recognition in reading and hearing **Müller, Bosch, Ernestus**	Context-aware XAI methods for joint human-AI problem solving **Finzel, Schmid**
3:40	Animacy cutweighs topichood when choosing referential expressions **Bader, Portele**	Patterns of representational change in problem solving of magic tricks **Danek, Williams, Wiley**
4:00	Unintentional response priming from verbal action–effect instructions **Damanskyy**	Where do problem spaces come from? On metaphors and representational change **Angerer**
4:20	Incremental negation processing with positive questions under discussion **Tsaregorodtseva, Albu, Kaup**	Discussion
4:40	**Coffee Break**	
5:00 – 6:20	POSTER SESSION	
6:30	SOCIETY MEMBER MEETING AULA	

Tuesday

8:30	**Registration**	
9:00	KEYNOTE 3 **Dedre Gentner** Analogy, Abstraction and Relational Knowledge AULA	
10:00	**Coffee Break**	
	SESSION 3 **SPACE, BODY & ACTION I** HÖRSAAL 1098	SYMPOSIUM 3 **CATEGORIZATION AND REPRESENTATION IN PERCEPTION, THINKING, AND ACTION** HÖRSAAL 1199
10:20	Gateway identity and spatial remapping in a combined grid and place cell attractor **Baumann, Mallot**	The relationship between colour perception and categorisation **Witzel**
10:40	Learning latent event codes for hierarchical prediction and generalization **Gumbsch, Martius, Butz**	Is there hierarchical generalization in action-effect learning? **Eichfelder, Franz, Janczyk**
11:00	Demands and potentials of different levels of neuro-cognitive models for human spatial cognition **Kuske, Röhrbein, Vitay, Ragni, Hamker**	On the role of categorization in evaluative conditioning: Category abstraction increases the generalization of acquired likes and dislikes **Reichmann, Hütter, Kaup, Ramscar**
11:20	A neural dynamic process model of scene representation, categorical visual search and scene grammar in natural scenes **Grieben, Schöner**	Towards unifying category learning and probability learning using the CAL framework of rule extrapolation and contextual modulation **Schlegelmilch, Wills, von Helversen**

11:40	A neural dynamic model of action parsing **Kang, Tekülve, Schöner**	An embodied neural process model grounds structured representations **Sabinasz, Schöner**
12:00	An embodied, perceptually grounded neural process model identifies analogies in accord with Structure Mapping Theory **Hesse, Sabinasz, Schöner**	Discussion
12:20	**Lunch Break**	
	SESSION 4 **SPACE, BODY & ACTION II** HÖRSAAL 1098	SYMPOSIUM 4 **DIGITAL CHILDHOOD: SOCIAL ROBOTS AND SOCIAL INTERACTIONS** HÖRSAAL 1199
1:40	The dynamics of body ownership: Extending the Bayesian causal inference of body ownership model across time **Schubert, Endres**	Talk, listen & keep me company: A mixed methods analysis of children's perspectives towards robot reading companions **Caruana, Moffat, Miguel-Blanco, Cross**
2:00	Inference of affordances and active motor control in simulated agents **Scholz, Gumbsch, Otte, Butz**	"It's not the robot who learns, it's me!" - considerations on the role of social robots in learning **Bruno**
2:20	The many faces of uncertainty in social interaction: a hierarchical model of metacognitive regulation in belief coordination **Kahl, Kopp**	Can a social robot advance children's long-term word learning of morphologically complex words by systematically varying the interaction? **Tolksdorf, Rohlfing**
2:40	**Coffee Break**	

	Session 5 **Foundations** Hörsaal 1098	Symposium 5 **Curiosity as a tool for information acquisition** Hörsaal 1199
3:00	Asymmetric Cross-Recurrence Plots for correlating time series of different length **Wallot, Felletti, Drews**	How curiosity affects learning and information seeking via the dopaminergic circuit **Gruber**
3:20	Deep neural networks as mechanistic explanations of object recognition - in search of the explanans **Grujicic**	A computational model of curiosity-based learning in infancy **Westermann, Twomey**
3:40	Making sense of the natural Environment **von der Malsburg, Grewe, Stadelmann**	Curiosity in social learning: infants' active information seeking from others in epistemic uncertainty **Bazhydai, Westerman**
4:00	Cognitive science between normativity and descriptivity **Strößner, Hahn**	How curiosity shapes early vocabularies **Ackermann**
4:20	Interfield integration in Cognitive Science **Poth**	Discussion
4:40	**Coffee Break**	
5:00 - 6:20	Panel Discussion Cognitive Science - Past, Present, and Future **Andrea Bender, Gregor Schöner, Dedre Gentner, Seana Coulson, Marcel Brass** **Aula**	
7:30	Conference Dinner **Restaurant Waldsee**	

Wednesday

8:00 **Registration**

9:00 PRESIDENTIAL ADDRESS & ANNOUNCEMENTS OF AWARDS
Best Publication Award Lecture
AULA

10:00 **Coffee Break**

	SESSION 6 **SOCIAL ISSUES IN LANGUAGE AND COGNITION** HÖRSAAL 1098	SYMPOSIUM 6 **COGNITION-CENTERED HUMAN-ROBOT INTERACTION** HÖRSAAL 1199	TUTORIAL 1 **NEURAL DYNAMICS FOR EMBODIED COGNITION** HÖRSAAL 1228
10:20	'Active attention guidance' as a basic scaffold for (social) cognition **Griem**	Robots as social agents: challenges and insights from social neuroscience **Wiese, Giebeler**	**Schöner, Tekülve**
10:40	Trading off subjective values and social expectations for interpersonal trust **Bellucci**	Bad robots? Humans rapidly attribute mental states during the perception of robot faces **Maier, Leonhardt, Rahman**	
11:00	Ambiguity remains a rare skill for learning about others even when stimuli carry social relevance **Achimova, Beukman**	Models of adaptive mentalizing for socially-aware HRI **Kopp**	
11:20	Gender associations elicited by masculine person-denoting nouns and indefinite pronouns **Bröder**	Integrating models of cognitive and physical human-robot interaction **Hao, Russwinkel, Haeufle, Beckerle**	

11:40	Gender priming non-linguistic stimuli: the effect of word and sentence primes on the perception of male, female, and gender-mixed face pairs **Kim, Öttl, Gygax, Behne, Hyönä, Gabriel**	What can cognitive modelling do for Human-Robot Interaction? **Heimisch**	TUTORIAL 1 cont.
12:00	The impact of social cognition on linguistic processing: Stereotypical information prevails over semantic encoding during reading **Mari, Müller**	Discussion	
12:20	**Lunch Break**		

1:40	KEYNOTE 4 **Seana Coulson** Understanding Words AULA
2:40	**Coffee Break**

	Session 7 **Knowledge** Hörsaal 1098	Symposium 7 **Sleeping – theoretical, developmental and applied perspectives on memory, dreams, and creativity** Hörsaal 1199	Tutorial 2 **Meta-Learned Models of Cognition** Hörsaal 1228
3:00	Reproducing children's category exemplars for fruit categories **León-Villagrá, Ehrlich, Lucas, Buchsbaum**	Using brain decoding to uncover traces of memory reactivation during rest and sleep **Kern, Gais, Feld**	**Marcel Binz**
3:20	Worked examples as application of analogical reasoning in Intelligent Tutoring and their effects on SQL competencies **Thaler, Mitrovic, Schmid**	The purpose of dreams **Anthes**	
3:40	Effects of inclusive dance training on body representations and body-related concepts in dancers with different bodily conditions **Bläsing**	Insightful problem solving in children's tool making: The role of napping **Gönül, Karabulut, Hohenberger**	
4:00	Open(ing) knowledge – concepts and discourses regarding Open Science **Dutschke, Lasch, Meier-Vieracker, Seemann, Scherbaum, Weigelt**	A digital intervention combining nightmare treatment and insomnia treatment **Lüth, Pipa, Gieselmann**	
4:20		Discussion	
4:40	Closing Remarks		

Posters

1	Signal enhancement or noise exclusion? Effect of temporal preparation on perceptual processes	**Janina Balke, Bettina Rolke, Verena C. Seibold**
2	Empirical Study on Auditory Perception in Autistic and Non-Autistic Hearers - Interplay between Prosody, Uncertainty, and Social Cognition	**Bellinghausen, Schröder, Rauh, Dahmen, Riedel, Fangmeier, Tebartz van Elst**
3	Relating ecological accounts of cognition with the reinforcement learning	**Olgierd Borowiecki**
4	Disentangling population stereotypes from individual differences in ACEs	**Francesca Capuano, Daniel Bub, Berry Claus, Barbara Kaup**
5	Competitive memory operations during discourse comprehension	**Susanne Dietrich, Verena C. Seibold, Bettina Rolke**
6	Relating Enactive Sense-Making to Psychological Meaning-Making	**Roy Dings, Caroline Stankozi**
7	On the processing of nominal compounds: the role of context and familiarity	**Gómez, Gamboa, Järvikivi, Allen**
8	The Effect of Feedback and Motivation on Time Estimation Performance	**Stine Hollah, Gerke Feindt, Sebastian Wallot**
9	Naming Together versus Naming Alone: A Mega-Analysis of Six Experiments on Joint Language Production	**Nora Holtz, Roger Hauber, Rasha Abdel Rahman, Anna K. Kuhlen**
10	“Flashing upon that inward eye” - A neurocognitive poetics approach to mental imagery	**Katharina Gloria Hugentobler, Jana Lüdtke**
11	Does deliberate prospection help students set better goals?	**Sarah Jähnichen, Felix Weber, Mike Prentice, Falk Lieder**
12	Role of Rehearsal in the Effects of Irrelevant Speech and Word Length	**Abdullah Jelelati, Larissa Leist, Thomas Lachmann, Maria Klatte**

13	Place-Name Categorisations Affect Wayfinding Behaviour	**Lilian LeVinh, Hanspeter Mallot**
14	A prototheory of climate cognition	**Lucas Lörch**
15	Using emotional word ratings to extrapolated norms for valence, arousal, imagebility and concreteness: The German list of extrapolated affective norms (G-LEAN)	**Jana Lüdtke, Katharina Gloria Hugentobler**
16	Core Hinges	**Jakob Ohlhorst**
17	A Study in …? German-English Interlingual Homographs and Their Organization in the Bilingual Brain – Evidence From Semantic Similarity	**Alina Palmetshofer, Carolin Dudschig, Fritz Günther, Barbara Kaup**
18	Discourse Illusions in L1 and L2	**Clare Patterson**
19	Using Cognitive Models and EEG Data To Investigate Spatial Cognition	**Kai Preuss, Nele Russwinkel**
20	Towards understanding spatial illusions in architecture - a pilot study exploring factors influencing illusive perspectives	**Vojtěch Rada, Julia Frankenstein, David Sedláček, Zdeněk Míkovec, Constantin A. Rothkopf**
21	The perception of Spanish vowels in adverse listening conditions: An investigation of adults with typical hearing, children with typical hearing, and children with hearing impairment	**Marcel Schlechtweg, Mark Gibson**
22	Investigating the effects of facial expressions and color cues on processing negated and affirmative sentences	**Emanuel Schütt, Merle Weicker, Carolin Dudschig**
23	Why do alerting signals increase congruency effects in the flanker task? An examination of boundary conditions	**Verena Carola Seibold**

24	Does contrastive attention guidance facilitate action recall? - An eye-tracking study.	**Amit Singh, Katharina J. Rohlfing**
25	Gaze behaviour during turn-taking in dyadic avatar-mediated conversations	**Malin Spaniol, Alicia Janz, Mathis Jording, Kai Vogeley**
26	Comparing sensory properties of words between English, Dutch, and Italian	**Annika Tjuka**
27	Text Comprehension: Mental Representation and Assessment	**Monika Tschense, Sebastian Wallot**
29	Structural Characteristics of Hierarchical Goal Systems from Online Field Studies	**Felix Weber**
30	Meet and Greet With Knowledge: A Study and Paradigm Investigating Text Comprehension Processes With Knowledge Models Generated by T-MITOCAR	**Tim Wilde, Anna Willisch, Wibke Maria Hachmann, Matthias Zaft, Pablo Pirnay-Dummer**
31	Logic and Psychology—Minding the Gap with Jean Piaget	**Mark Anthony Winstanley**

Abstracts

Invited Keynote Presentations

Free Will: An Empirical Perspective

Marcel Brass
Humboldt University, Berlin, Germany

The question whether free will exists or not has a century old history in philosophy. Only a few decades ago, however, neuroscientists and psychologists have started to investigate this question empirically. Based on neuroscientific findings, prominent researchers have claimed that free will is an illusion. Such claims had a strong impact on the free will debate both in psychology and philosophy. Moreover, these claims have found their way into the public media, potentially affecting free will beliefs of people outside academia. In my talk, I will first discuss the validity of the claim that neuroscience has disproven free will. Furthermore, I will address the question if it matters whether people believe in free will or not.

The Neurocomputation of Sentence Meaning

Matthew Crocker
Saarland University, Saarbrücken, Germany

I will outline recent results from my lab supporting the view that event-related brain potentials directly index two core mechanisms of sentence comprehension: association-driven retrieval of each word from semantic memory (N400) and expectation-based semantic integration with the unfolding sentence meaning (P600). I will then present a neurocomputational model of language comprehension which directly instantiates these two systems, including explicit linking functions to both behavioural and neurophysiological comprehension measures.

Analogy, Abstraction and Relational Knowledge

Dedre Gentner

Northwestern University, Evanston, IL, USA

A prominent stance in recent Cognitive Science is that human concepts are naturally embodied—concrete and contextually situated. Whether we also possess abstract concepts has been controversial. I will argue that we form abstract relational concepts from embodied experience and that we do so via analogical comparison—specifically, via structural alignment and mapping. Although structure-mapping ability is present early in infancy, analogical ability is vastly increased by acquiring language and other symbol systems. In particular, metaphoric language is a route to new abstractions both in individual learning and in language evolution. I'll present evidence from children and adults to support these claims.

Understanding Words

Seana Coulson

University of California, San Diego, CA, USA

What is it to understand the meaning of a word? Distributional semantics provides one sort of answer to this question. Word meanings are multidimensional vectors induced from the linguistic contexts in which they occur. Embodied meaning provides a somewhat different answer: word meanings include (in part) simulations that recruit sensorimotor activations. I discuss empirical data from my lab that bears on each of these proposals, paying special attention to the N400, an ERP component conventionally associated with semantic retrieval. Finally, I consider whether an unholy alliance between these proposals is possible or desirable.

Symposia

Symposium 1

Symposium 1: **“*Computational modelling of the Active Self*”**

Organizers: Prof. Nele Rußwinkel (TU Berlin) & Prof. Stefan Kopp (University of Bielefeld)

The Active Self is commonly conceived as a concept related to one’s phenomenal experience when acting in the here and now, and to the question of how we perceive ourselves to be in a particular situation. Research into this topic has gained considerable momentum and involves cognitive and behavioural science, psychology, neuroscience, robotics and other disciplines. What mechanisms are responsible for the plasticity of our self-representation and sensitivity to immediate experience? How are experiences of agency and ownership during motor action integrated into a minimal concept of an Active Self? How does the Active Self interact with and affect other cognitive processes? These are general questions that research in this area poses. Besides theoretical discussion and approaches to measure it empirically, computational modelling of the cognitive and embodied processes underlying action control, motor learning, or multi-sensory perception has started to shed additional light on these questions. And, such modelling approaches open up new possibilities for interdisciplinary exchange about concepts and theories as every modelled aspect or mechanism has a specific, testable instantiation.

In this symposium, different aspects of how computational models can help to elevate our understanding of the Active Self will be presented and discussed. Topics covered may include how controllability can be perceived and achieved, how a self can be maintained under external disturbances, how a neurological plausible model can explain effects of the active self, how symbolic and sub-symbolic implementations can be integrated, and how the self is involved in anticipating own but also others’ actions. The different talks will be presented, each addressing the following questions for its particular modelling perspective:

- What aspect(s) of the active self is addressed in the modelling approach, and what do we learn from that approach?
- Why was that particular modelling method chosen to address this problem?
- How can other researches benefit from this model?
- Under which circumstances is it necessary or beneficial, to relate different modelling approaches to each other or integrate them?

The confirmed list of speakers includes Prof. Verena Haffner (HU Berlin), Prof. Gregor Schöner (Ruhr-Universität Bochum), Prof. Martin Butz (Uni. Tübingen) and a joint talk by the organizers. Together they will represent an overview of research in the current DFG priority program on the Active Self. A general discussion at the end of the symposium will identify links and provide a broader thematic elaboration of overarching aspects.

Computational cognitive modeling of the predictive active self in situated action

Nele Russwinkel [a] & Stefan Kopp [b]

[a] Cognitive Modeling in dynamic Human-Machine Systems, Department of Psychology and Ergonomics, Technische Universität Berlin, Germany

[b] Social Cognitive Systems Group, CITEC, Faculty of Technology, Bielefeld University, Germany

Autonomous embodied agents in an uncertain real-world environment need to continuously adapt their situated action. One important mechanism for this is a sense of control (SoC), i.e. the assessment of how one is in control of one's own actions and their effects. We present work towards investigating how the SoC arises from and affects the cognitive and sensorimotor processes underlying the control of situated actions (Kahl, et al., 2021). To that end, we combine experiments with human participants with computational cognitive modeling and simulation-based studies. We will present work in which we have (1) implemented a cognitive architecture for prediction-based action control by integrating an established model of cognitive processes (ACT-R) with a probabilistic model of predictive sensorimotor processes, (2) modeled the sensorimotor and cognitive components of SoC as well as their interplay in the top-down/bottom-up processes underlying prediction- based action control, (3) designed and applied a task scenario for experimental data collection and model testing through simulation. Further, we will outline next steps investigating the extent to which the active self is imparted to task-specific representations and thus influences the cognitive processes that control how concurrent tasks are managed.

Keywords: Cognitive Modelling, Active Self, Situated cognition, Action Control

References

Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. (2021). Towards autonomous artificial agents with an active self: modeling sense of control in situated action. Cognitive Systems Research, 72, 50-62. https://doi.org/10.1016/j.cogsys.2021.11.005

Tool-use and agency in artificial agents

Verena V. Hafner
Adaptive Systems Group, Department of Computer Science
Humboldt-Universität zu Berlin, Germany

Sense of agency and sense of body ownership are important prerequisites for a minimal self in humans. When we look into the prerequisites for such a minimal self in artificial agents, a computational sense of agency can be created by predictive internal models that are able to predict the sensory consequences of a particular motor action of the agent based on previous experience (Hafner et al., 2020). The predictive models need to be adapted to the current context of the agent and its environment. Here, we will present a variety of models (Ciria et al., 2021) and discuss how they can be used for decision making and tool-use.

Sense of Agency

A sense of agency in humans is the experience of perceiving actions as self-generated by anticipating the sensory consequences of a performed motor action based on previous experience. A simple example would be kicking a ball or moving a glass on a table closer to a bottle.

In artificial agents, this can be realized by using predictive models (Ciria et al., 2021; Friston, 2018) to predict the sensory consequences of a performed action of an artificial agent. The predictive models can be learned from data that were created during previous exploration processes (Schillaci et al., 2016; Hafner et al., 2020).

Tool-Use

It has been proposed that tools can be included in our body schemas (Maravita & Iriki, 2004). We could then also perceive actions with a tool as self-generated the same as without a tool.

When using predictive models in artificial agents, the models could also be trained by exploration using tools. We will show based on a simple decision task with and without tools in robots, that predictive models can help to choose the correct tool for a given task (Schillaci et al., 2012).

Keywords: Artificial Self, agency, body ownership, tool-use.

References

Ciria, A., Schillaci, G., Pezzulo, G., Hafner, V. V., & Lara, B. (2021). Predictive Processing in Cognitive Robotics: A Review. *Neural Computation*, *33*(5), 1402-1432. Retrieved from https://doi.org/10.1162/neco_a_01383 doi: 10.1162/neco_a_01383

Friston, K. (2018, 07). Does predictive coding have a future? *Nature Neuroscience*, *21*. doi: 10.1038/s41593-018-0200-7

Hafner, V. V., Loviken, P., Pico Villalpando, A., & Schillaci, G. (2020). Prerequisites for an artificial self. *Frontiers in Neurorobotics*, *14*. Retrieved from https://www.frontiersin.org/article/10.3389/fnbot.2020.00005 doi: 10.3389/fnbot.2020.00005

Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in cognitive sciences*, *8*(2), 79–86.

Schillaci, G., Hafner, V. V., & Lara, B. (2012). Coupled inverse-forward models for action execution leading to tool-use in a humanoid robot. In *Proceedings of the seventh annual acm/ieee international conference on human-robot interaction* (p. 231?232). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2157689.2157770 doi: 10.1145/2157689.2157770

Schillaci, G., Hafner, V. V., & Lara, B. (2016). Exploration behaviors, body representations, and simulation processes for the development of cognition in artificial agents. *Frontiers in Robotics and AI*, *3*. Retrieved from https://www.frontiersin.org/article/10.3389/frobt.2016.00039 doi: 10.3389/frobt.2016.00039

A neural dynamic account of intentionality as the basis of an active self

Gregor Schöner and Jan Tekülve
Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany

How are sequences of motor behaviors coordinated with processes of active perception to enable an agent to direct actions at objects in its environment in order to achieve desired outcomes? We postulate that moving beyond the simplest reflex chains or fixed action patterns, an account for this competence of organism can be provided only together with a neural account of intentionality, that is, of how mental states can be about relevant states of affairs in the world. This is because planning ahead, acquiring and using knowledge to achieve goals all require such mental states. Intentionality, we postulate, is the basis of the active self.

Neural process accounts of intentionality must link intentional states to sensory processes and motor systems. Such links require that intentional states posses stability, that is, resistance to change under varying conditions. Neural processes that ensure the match between states of the world and states of the mind, the conditions of satisfaction are critical to dissolving that stability within sequences of intentional states needed to achieve goals.

To make these ideas concrete, we build a neural dynamic model of intentionality in a a highly simplistic toy scenario in which a robotic agent learns to achieve goals. The agent is a vehicle with a robotic arm that perceives the world through a vision sensor and has a set of sensors to perceive its body state. The environment consists of cubes that are either paint buckets or function as canvases to be painted on. The agent moves through space, pointing its arm at objects to pick up or deposit paint. It learns color contingencies (which color results when applying a given coat of paint to a canvas of a given color) and uses that knowledge to achieve goals, the simple desire to see objects of a particular color.

Homeostasis Drives the Active Self while Generative Models Constitute It

Martin V. Butz, martin.butz@uni-tuebingen.de, Neuro-Cognitive Modeling Group, Department of Computer Science and Department of Psychology, Faculty of Science, Sand 14, University of Tübingen, Germany

My goal in this abstract is to integrate two research directions in cognitive science: the principle of free energy minimization and event-predictive cognition, and to discuss the implications for theories of an active self. Predictive encoding and the free energy minimization principles have developed from an idea to a fundamental brain theory over the last twenty years (Rao & Ballard, 1999; Friston, 2009). Their roots reach back into the 19th century and the ideomotor principle (Herbart, 1825; Stock & Stock, 2004), which emphasizes that initially reactive behavior soon becomes goal-directed and self-motivated: desired goals trigger the associated behavior, potentially modulated by the context. Correspondingly, the free energy principle formalizes how generative models may co-develop with the efficiency, flexibility, and adaptivity of goal-directed behavior by both minimizing expected uncertainty and maximizing expected homeostasis. Thereby, the developing probabilistic model learns to generate hypothesis about causes, which determine our sensory perception, and forces, which bring about change. It naturally focuses on those aspects of the environment that appear action-relevant, that is, causally relevant for achieving particular goals.

At this point, it may still be unclear where intentions and desired internal states that determine homeostasis come from. These states may include genetically predetermined, but also environmentally and culturally-shaped homeostatic states. The active inference formalism (Friston et al., 2015) suggests that our behavior is dynamically shaped by our needs to avoid uncertainty and pursue homeostasis. Thereby, homeostasis is quantified by the negative anticipated divergence of action-conditioned state predictions from desired ones. For humans these states do not only encode bodily but also social needs. The latter are driven by cultural practices and a desire to find meaningful social roles for our own selves (Tomasello, 2019).

Meanwhile, event-predictive inductive learning and processing biases shape the development of mental generative models (Butz, 2016; Butz et al., 2021). These biases foster the emergence of compact compositional encodings of our experiences, including causal dependencies between them. Combined with the tendency to develop event-predictive models, the active inference process enables us to consider events that lie further into the future and may concern alternative or counterfactual developments. Moreover, it becomes possible to recombine courses of events in an innovative compositional manner. Our own recent computational modeling work has shown that these mechanisms can model the development of anticipatory eye saccades in infants (Gumbsch et al., 2021). A neural implementation furthermore shows that stable hidden context-encodings foster generalization, enable one-shot learning, and offer explainable, conceptual encoding structures (Gumbsch et al., 2021).

Finally, I consider how our brains' limited cognitive resources influence the developing active self. Cognitive resources are needed to draw inferences within our event-predictive generative models (Butz, 2022; Zénon et al., 2019). These limits constrain our rationality (Lieder & Griffiths, 2020). Moreover, to save resources, our brains tend to build habitual behavioral patterns that have previously generated advantageous outcomes (Herbort & Butz, 2011; O'Doherty et al., 2017). These habitual behaviors are energy efficient but come at the cost of sometimes hindering rational and educated decision making and belief inference. As a result, resource limitations and habitual behaviors further bias our active selves.

I thus maintain that our active selves are constituted by generative event-predictive models and are shaped and driven by our homeostatic makeup, constrained by the cognitive processing capacities available to us.

Keywords: predictive coding, active inference, homeostasis, event-predictive cognition, goal-directed behavior, habitual behavior, concept learning

References

Butz, M. V. (2016). Towards a unified sub-symbolic computational theory of cognition. *Frontiers in Psychology*, *7*(925). doi: 10.3389/fpsyg.2016.00925

Butz, M. V. (2022). Resourceful event-predictive inference: The nature of cognitive effort. *Frontiers in Psychology*, *13*. doi: 10.3389/fpsyg.2022.867328

Butz, M. V., Achimova, A., Bilkey, D., & Knott, A. (2021). Event-predictive cognition: A root for conceptual human thought. *Topics in Cognitive Science*, *13*, 10-24. doi: 10.1111/tops.12522

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293 - 301. doi: 10.1016/j.tics.2009.04.005

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., FitzGerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*, 187-214. doi: 10.1080/17588928.2015.1020053

Gumbsch, C., Adam, M., Elsner, B., & Butz, M. V. (2021). Emergent goal-anticipatory gaze in infants via event-predictive learning and inference. *Cognitive Science*, *45*(e13016). doi: 10.1111/cogs.13016

Herbart, J. F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil [Psychology as a science newly grounded on experience, metaphysics, and mathematics. second part: Analytics]*. Königsberg, Germany: August Wilhem Unzer.

Herbort, O., & Butz, M. V. (2011). Habitual and goal-directed factors in (everyday) object handling. *Experimental Brain Research*, *213*, 371-382. doi: 10.1007/s00221-011-2787-8

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. doi: 10.1017/S0140525X1900061X

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annu. Rev. Psychol.*, *68*(1), 73–100. doi: 10.1146/annurev-psych-010416-044216

Rao, R. P., & Ballard, D. H. (1999, January). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. doi: 10.1038/4580

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research*, *68*, 176-188. doi: 10.1007/s00426-003-0154-5

Tomasello, M. (2019). *Becoming human:a theory of ontogeny*. Harvard University Press.

Zénon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, *123*, 5-18. doi: 10.1016/j.neuropsychologia.2018.09.013

Symposium 2

Flexible Representations for Human(-like) Problem Solving

Benjamin Angerer
Thea Berens
Amory Danek
Bettina Finzel

The challenges of problem solving do not exclusively lie in how to perform heuristic search, but they begin with how we understand a given task: How to cognitively represent the task domain, its components, and the (sub-)goals an agent tries to achieve can determine how quickly progress towards a solution is being made, whether advanced strategies can be discovered, and sometimes even whether a solution can be found at all.

Especially for more complex task domains, there can be a wide variety of potential representations, and it might even be beneficial to make simultaneous use of several ones, or to change them during the problem solving process. While this challenge of constructing and changing representations has been acknowledged early on in problem solving research, for the most part it has been sidestepped by focussing on simple, well-defined problems whose representation is almost fully determined by the task instructions. Thus, the established theory of problem solving as heuristic search in problem spaces has little to say about these issues.

However, over time there have been many developments in related fields which might play a role in addressing this impasse, e.g. the mechanisms of analogy-making, metaphor use, and explanation, the contribution of affective, bodily and environmental factors to cognitive processes, and the development of more complex and ecologically valid experimental tasks eliciting a broader range of behaviour. Yet, so far few of these research fields have integrated their insights into a common problem solving theory, leaving the issues of representation still largely unaddressed.

In this symposium, we bring together researchers working on insight, metaphors, strategy discovery, and explainable AI in order to reflect on the current and future state of problem solving theory and research with respect to understanding the mechanisms of flexible and dynamic representations in humans, artificial systems, and their interaction.

Contributed talks

- Thea Behrens (Models of Higher Cognition Group, Technical University of Darmstadt)
  Strategies and Representations for Solving Sudokus

- Bettina Finzel (Cognitive Systems Group, University of Bamberg)
  Context-Aware XAI Methods for Joint Human-AI Problem Solving

- Amory Danek (Experimental and Theoretical Psychology, University of Heidelberg)
  Patterns of representational change in problem solving of magic tricks

- Benjamin Angerer (virtUOS, University of Osnabrück)
  Where do problem spaces come from? On metaphors and representational change

Strategies and Representations for Solving Sudokus

Thea Behrens, Frank Jäkel
Institute of Psychology, Technical University Darmstadt
Centre for Cognitive Science, Technical University Darmstadt

Sudoku Puzzles

Sudoku is a number placement puzzle which can be described as a constraint satisfaction problem (CSP). A standard Sudoku is a partially filled 9-by-9 grid, which is additionally separated into nine 3-by-3 boxes. The empty cells have to be filled by the player and there is just one correct solution. Each number has to occur exactly once per row, column and box. These rules are just a statement of the constraints, they do not directly provide an algorithm for solving a Sudoku. To successfully fill a Sudoku, a player thus has to come up with strategies to find enough constraints to be certain what number to put into which cell.

Possible Problem Representations

There are at least two quite distinct representations one can form of the problem and use as a basis for finding a solution to a Sudoku puzzle. One representation uses the empty cells of the Sudoku as variables and the numbers from 1 to 9 as potential values for the variables. In this case it is natural to look for the most constrained cells: those where in the same row, column or box (or a combination of these) many numbers are already assigned and can thus be excluded for the cell under consideration. If all values except one can be excluded, the remaining value can be assigned to the cell. The other representation uses nine instances of each number as variables and the cells as values to assign to them. One can, for example, look for *where in a box* one can place a given number. Some locations can be excluded for a number either because they are filled or because the number is already present in the row or column crossing the cell. When only one cell in a unit remains as possible candidate for a number, it can be filled in. The two representations afford different solution tactics and they are suitable for different situations.

Experimental Evidence for Flexible Switching

The existence of these two distinct representations of the problem make Sudoku a great domain to study flexibility in switching between representations and tactics in problem solving. I will present data from three experiments (Behrens et al., 2022). In a think-aloud study participants used both representations and had personal preferences for one or the other. Response times in two follow-up experiments indicate that participants can be biased towards either representation by task instructions. We argue that previous research often used biasing task designs and therefore underestimated participants' flexibility. Furthermore, our experiments demonstrate that if no solution can be found within one problem representation, participants are able to switch to the other. We developed a process model that implements tactics based on both representations and matches the patterns of response times of all conditions of our experiments. Fitting the model to the response time data shows us in greater detail what the most likely preferred representation of each participant is.

Keywords: human problem solving, problem representation, strategy selection

References

Behrens, T., Räuker, M., Kalbfleisch, M., & Jäkel, F. (2022). Flexible use of tactics in Sudoku. *Thinking & Reasoning*, 1–43. doi: https://doi.org/10.1080/13546783.2022.2091040

Context-aware XAI Methods for Joint Human-AI Problem Solving

Bettina Finzel, Ute Schmid
Cognitive Systems, University of Bamberg
An der Weberei 5, 96047 Bamberg, Germany

Explainable artificial intelligence (XAI) helps human experts and end-users solve complex problems as it gives insights into decisions made by supporting AI systems (Miller, 2019). However, current approaches primarily generate static and unimodal explanations, for example, in the form of highlighting. They contrast with the richness and flexibility of human explanations, particularly with the adaptation of explanations to the context of problem-solving. Therefore, methods are needed that consider explaining as the recipient's learning process and that produce adaptable and multimodal explanations (Finzel et al., 2021; Schmid & Finzel, 2020; Rabold et al., 2019, 2021). In this talk, we will present visual and verbal explanation methods suitable for implementing such a multimodal explanation process, including contrastive, prototype-based, dialogical, and interactive approaches. This talk will also introduce which aspects should be considered when adapting explanations to the recipient's context. According to Abowd et al.'s concise definition of context (Abowd et al., 1999), we are mainly guided by the dimensions of location (where?), identity (who?), activity (what for?), and time (when?). That is, one can consider explanations as entities that have locations in an explanatory space from which one chooses appropriate representations (e.g., global or local, symbolic or sub-symbolic representations). Furthermore, it is essential to consider who is being explained to (Kulesza et al., 2013). What prior knowledge does a person have? The purpose of the explanation and the recipient's goal must also be taken into account. Last but not least, context-aware approaches should also consider when the human decision-maker needs the information. We show that adaptable and multimodal approaches are suitable to represent different dimensions of context in explanations and to improve joint human-AI problem-solving.

Keywords: context-aware XAI, human-AI partnership, multimodal explanations, explanation as a process

References

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999). Towards a Better Understanding of Context and Context-Awareness. In *International Symposium on Handheld and Ubiquitous Computing* (pp. 304–307).

Finzel, B., Tafler, D. E., Scheele, S., & Schmid, U. (2021). Explanation as a Process: User-centric Construction of Multi-level and Multi-modal Explanations. In *German Conference on Artificial Intelligence (Künstliche Intelligenz)* (pp. 80–94).

Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., & Wong, W.-K. (2013). Too Much, Too Little, or Just Right? Ways Explanations Impact End Users' Mental Models. In *2013 ieee Symposium on Visual Languages and Human Centric Computing* (pp. 3–10).

Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the Social Sciences. *Artificial Intelligence*, *267*, 1–38.

Rabold, J., Deininger, H., Siebers, M., & Schmid, U. (2019). Enriching Visual with Verbal Explanations for Relational Concepts – Combining LIME with Aleph. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 180–192).

Rabold, J., Siebers, M., & Schmid, U. (2021). Generating Contrastive Explanations for Inductive Logic Programming Based on a Near Miss Approach. *Machine Learning*, 1–22.

Schmid, U., & Finzel, B. (2020). Mutual Explanations for Cooperative Decision Making in Medicine. *KI-Künstliche Intelligenz*, *34*(2), 227–233.

Patterns of representational change in problem solving of magic tricks

Amory H. Danek[1], Joshua Williams[2], Jennifer Wiley[2]

[1] Psychologisches Institut, Universität Heidelberg, Heidelberg, Germany

[2] Department of Psychology, University of Illinois at Chicago, Chicago, USA

Background:

Hallmarks of insightful problem solving are thought to be suddenness in the emergence of a solution (Metcalfe & Wiebe, 1987), an underlying representational change (Knoblich et al., 1999; Ohlsson, 1992), and the subjective Aha! experience (Kaplan & Simon, 1990). However, little work has demonstrated a connection among these cognitive and affective solution features, although this is theoretically assumed (Dominowski & Dallob, 1995).

Method:

In the present study, we used a problem set consisting of 18 magic tricks, combined with a self-report measure of representational change to track participants' problem solving process. A sample of 30 students attempted to detect the secret method behind each trick, after watching short video clips of the tricks three times. Trial-wise, ratings of Aha! experiences were obtained upon solving. In addition, we used the following procedure to assess initial problem representations and to track changes in representations during solution: After each viewing of a trick, solvers rated a list of action verbs for how important they seemed for solution, in order to capture the solution that was currently being considered by the solver (as done by Ash & Wiley, 2008; Cushen & Wiley, 2012; Danek et al., 2020).

Results:

Initial ratings showed that solvers mostly started with incorrect problem representations. Changes in ratings across the three rating time points that went from selecting inappropriate verbs to selecting the verb that matched the actual solution provided evidence of representational change. The resulting patterns of change in solvers' mental problem representations were categorized by independent raters as sudden or incremental (plus some further categories). We hypothesized that sudden-change patterns ("insight-like") should be more likely than incremental ones to be connected to Aha! experiences. Indeed, solutions following sudden patterns received higher Aha! ratings than solutions following incremental patterns. These results offer support for prior assumptions of a close relationship between the three main aspects of insight.

Keywords: Insight, problem solving, representational change, restructuring, solution patterns.

References

Ash, I. K., & Wiley, J. (2008). Hindsight bias in insight and mathematical problem solving: Evidence of different reconstruction mechanisms for metacognitive versus situational judgments. *Memory & Cognition, 36*(4), 822–837.

Cushen, P. J., & Wiley, J. (2012). Cues to solution, restructuring patterns, and reports of insight in creative problem solving. *Consciousness and Cognition, 21*(3), 1166–1175.

Danek, A. H., & Wiley, J. (2020). What causes the insight memory advantage? *Cognition, 205,* 104411.

Dominowski, R. L., & Dallob, P. (1995). Insight and problem solving. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 33–62). Cambridge, MA: MIT Press.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology,* 22(3), 374–419.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1534–1555.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition, 15, 238–246.*

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. T. Keane & K. J. Gilhooly (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–44). Harvester-Wheatsheaf.

Where do problem spaces come from? On metaphors and representational change

Benjamin Angerer*
VirtUOS, University of Osnabrück
Institute of Philosophy, University of Osnabrück

50 years after the influential formulation of the theory of human problem solving as heuristic search in problem spaces (Newell & Simon, 1972), we still possess limited knowledge about how problem representations are generated and – if necessary – changed over time (Ohlsson, 2012). To begin addressing this issue, a study was conducted in which finding and refining an adequate problem representation was the main challenge of the task.

In this exploratory case study, it was investigated how pairs of participants acquaint themselves with a complex task domain over the course of several days. Specifically, they had to work on a complex spatial transformation and problem solving task in the domain of iterated mental paper folding (Angerer & Schreiber, 2019). In this task, participants have to understand the geometry of edges which occurs when repeatedly mentally folding a sheet of paper in alternating directions without being allowed the use of any external aids. While the *activity* of cross-folding sheets is familiar, the consequences of what exactly happens to the sheet's geometry are more difficult to understand. Faced with the difficulty of mentally handling increasingly complex folds (Fig. 1), participants are forced to look for regularities in the folding process, such as repeating configurations of edges, common transformations between folds, etc. which help them to represent folds more efficiently. In a qualitative analysis of video recordings of the participants' behaviour, the development of their conceptualisation of the task domain was traced over the course of the study, focussing especially on the use of gesture and the spontaneous occurrence and use of idiosyncratic metaphors in the construction of new representations (Fig. 2).

Based on these observations, we suggest that change is driven by a trade-off between the currently perceived task demands and the expressive limitations of the representational substrates in question. Iconic representations, such as gestures, are useful in representing simple objects, but direct resemblance soon encounters limits. Metonymic gestures, iconically resembling parts of an object but referring to the whole, can extend this scope. By thus omitting aspects of the task domain, those partial resemblances can in turn feed associative memory processes generating metaphors which bear no actual similarity to domain objects but to, for instance, such metonymic gestures. Finally, elaborating on such metaphors can open up further representational possibilities.

Keywords: metaphor, problem solving, representational change, gestures, case study

*Email: benjamin.angerer@uos.de

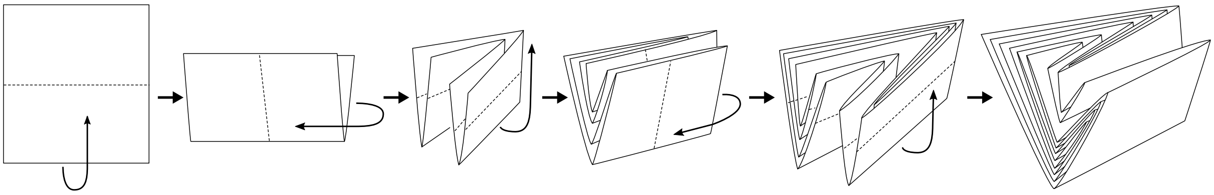


Figure 1: Cross-folding a sheet five times by alternating between two perpendicular folding directions. While the geometry of edges produced by iterated cross-folding is governed by a simple set of recursive rules, this is not easy to see.

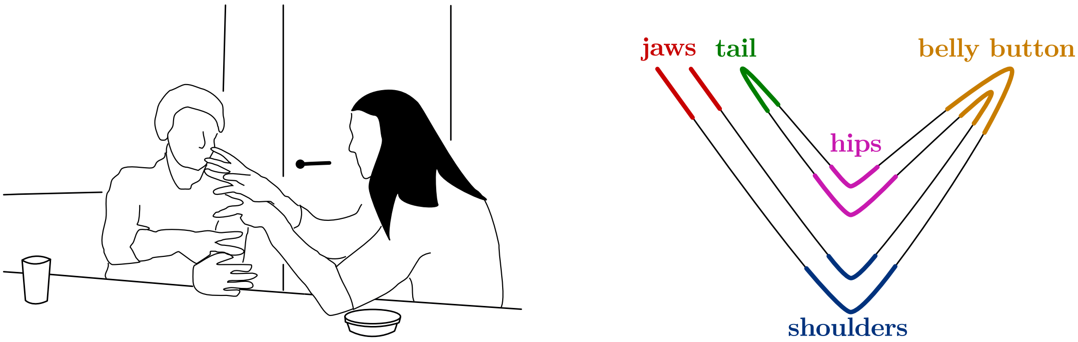


Figure 2: Left: Reasoning about a cross-fold with a joint gesture in which fingers correspond to individual edges, *Right*: Mnemonic landmarks on a "doubly-folded crocodile" (a single edge which has been folded 3 times): An idiosyncratic metaphor derived from blending generic anatomical knowledge and the shape-based likening of a single-folded edge with a crocodile's mouth.

References

Angerer, B., & Schreiber, C. (2019). Representational dynamics in the domain of iterated mental paper folding. *Cognitive Systems Research*, *54*, 217–231.

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S. (2012). The Problems with Problem Solving: Reflections on the Rise, Current Status, and Possible Future of a Cognitive Research Paradigm. *Journal of Problem Solving*, *5*(1), 101–128.

Symposium 3

Symposium 3:

Categorization and Representation in Perception, Thinking, and Action

Karin M. Bausenhart & Kathrin Reichmann

University of Tübingen

Categorization is a core function of human cognition (for overviews, see, e.g., Cohen & Lefebvre, 2005; Goldstone, Kersten, & Carvalho, 2013; Kelter & Kaup, 2019). Via categorization, distinct features and interrelations can be grouped together into broader, coherent units of knowledge (i.e., mental concepts), allowing for the organization of objects, events and actions in the mind. The process of categorization entails abstraction, that is, removal of irrelevant features, as well as prediction, namely, making inferences based on category membership. For instance, we may categorize a small, red, roundish object as an apple, and infer from conceptual knowledge that it must be edible and of a certain weight. This, in turn, would enable programming of the appropriate motor parameters such as grip aperture and the force needed to pick it up and take a bite. Thus, categorization processes are relevant for various domains of cognitive processing, as for example the interpretation and understanding of sensory inputs or the execution of actions. Unsurprisingly, a plethora of research paradigms in different cognitive domains have been developed to understand the principles underlying categorization, but it is unclear how these approaches relate to each other. Consequently, it remains an open question whether the various categorization processes can be understood within a universal framework, rather than requiring a unique characterization for different processing domains. The present symposium invites researchers studying categorization from different perspectives, aiming to improve our understanding of the role and impact of category knowledge at various stages of information processing across perception, thinking, and motor processing. Specifically, we will include research looking at categorization in colour perception (Witzel), probability learning (Schlegelmilch et al.), attitude acquisition (Reichmann et al.), and action-effect learning (Eichfelder et al.). Bringing together these different research perspectives may highlight commonalities and differences in the underlying principles and eventually help to gain a more unified understanding of categorization processes and their role in the various domains of human information processing.

References

Cohen, H., & Lefebvre, C. (Eds.). (2005). *Handbook of categorization in cognitive science*. Amsterdam: Elsevier.

Goldstone, R. L., Kersten, A., & Carvalho, P. F. (2013). Concepts and categorization. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 4, pp. 599–621).
Hoboken, NJ: Wiley.

Kelter, S. & Kaup, B. (2012). Conceptual knowledge, categorization, and meaning. In C. Maienborn, K. Heusinger & P. Portner (Ed.), *Semantics (Vol. 3): An international handbook of natural language meaning* (pp. 2775-2804). Berlin, Boston: De Gruyter Mouton.

Vorträge

The relationship between colour perception and categorisation

Christoph Witzel
University of Southampton

Towards unifying category learning and probability learning using the CAL framework of rule extrapolation and contextual modulation

René Schlegelmilch[1], Andy J. Wills[2], & Bettina von Helversen[1]
[1] University of Bremen
[2] University of Plymouth

On the "boxes" in "putting someone into boxes": category abstraction increases the generalization of acquired likes and dislikes

Kathrin Reichmann, Barbara Kaup, Michael Ramscar, & Mandy Hütter
University of Tübingen

No evidence for hierarchical generalization in action-effect learning

Lea Eichfelder[1], Volker Franz[2], & Markus Janczyk[1]:
[1] University of Bremen, Germany
[2] University of Tübingen, Germany

The Relationship Between Colour Perception and Categorisation

Christoph Witzel, University of Southampton, cwitzel@daad-alumni.de

This contribution to the symposium discusses our recent neuropsychological and developmental studies on colour categorisation and our ecological account of colour categorisation. Colour categorisation has been a prime example to investigate the relationship between perception and language (for review see, e.g., Witzel & Gegenfurtner, 2018; Witzel, 2018). There is an explanatory gap between the way we perceive colours continuously in terms of hue, saturation and lightness, and the way we categorise colours when we use basic colours terms, such as red, purple, and pink (cf. Figure 1).

Background: Classical Approaches

For many decades, the field has debated whether colour perception determines colour categorisation, or, conversely, whether language acquisition influences the way we see colours. Studies on categorisation within and across languages and on categorical perception were key approaches to establish a relationship between perception and language. However, evidence for a linguistic origin remained ambiguous, suggesting some language-specific influences besides universal determinants of colour categories (for review, e.g., Witzel, 2018). It also turned out that colour perception per se is not categorical. There may be high-level, cognitive effects of categorisation on performance in perceptual and memory tasks, but these effects are rather small and vulnerable to task and measurement conditions (for review, e.g., Siuda-Krzywicka, Boros, et al., 2019; Witzel & Gegenfurtner, 2018).

Categorisation independent of language and perception

Recently, we proposed that colour categorisation might be a phenomenon separate from both, colour perception and lexical colour naming. This idea is supported by neuropsychological evidence, according to which colour categorisation may be preserved when colour naming is damaged due to a left-hemisphere lesion (Siuda-Krzywicka, Witzel, et al., 2019). In another study measuring resting state fMRI, we identified separate functional brain networks for colour naming and categorisation (Siuda-Krzywicka, 2020). Categorisation might rely on an ability to generalise across instances that is neither specific to colour perception nor to lexical naming. Learning to recognise common patterns across varying instances might also be the basis for the relationship between the development of colour constancy and colour categorisation during colour term acquisition (Rogers et al., 2020; Witzel et al., 2021).

Ecological Origin of Colour Categories

Rather than being tied to either perception or language, we suggest colour categories might have an ecological origin (e.g., Siuda-Krzywicka, Boros, et al., 2019; Witzel, 2018; Witzel & Gegenfurtner, 2018). This approach focuses on the functional role of colour categories in recognising and communicating objects, materials, or other scene characteristics in the environment. Through their experience, observers learn which colour differences are important for communicating about the visual environment, and which are not and can be ignored by grouping colours together. For example, the colour term "green" might have developed to describe the colour of plants. According to our ongoing research, speakers of many fundamentally different languages associate plants with their colour term for "green". The large variety of leave colours would also explain why green is such a large category compared to others (cf. Figure 1). The fact that there are stable and variable properties in the visual environment may explain universal motifs and variation across languages, respectively.

Keywords: Categorisation, Colour Naming, Colour Perception, Perception and Language, Sapir-Whorf-Hypothesis

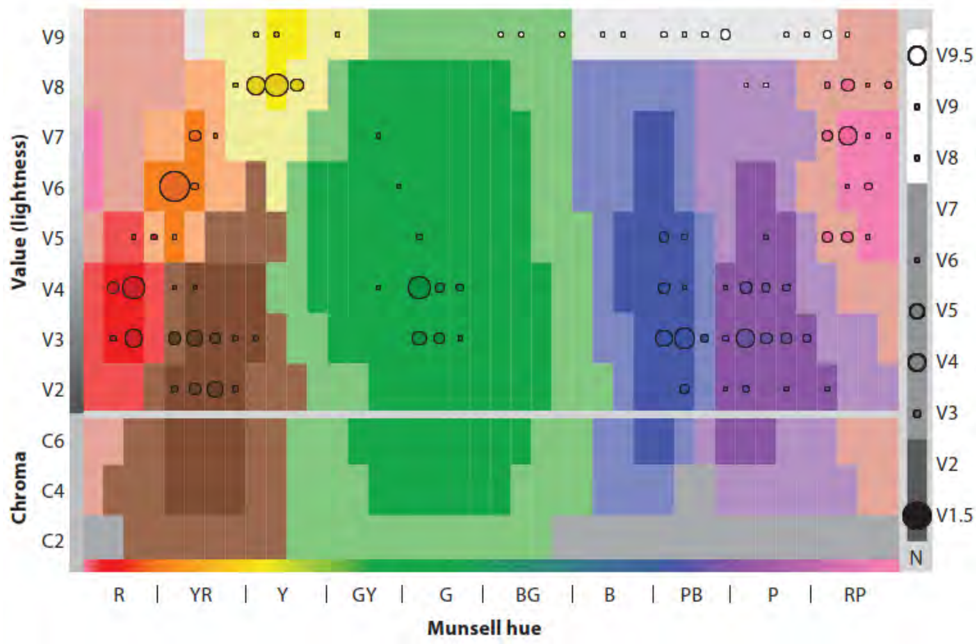


Figure 1. Categorisation of Perceptual Colour Space in Colour Naming. Rectangles (background) correspond to colour chips. The colours of the rectangles indicate the mode colour terms chosen by observers. Colour chips are arranged by their hue (x-axis) and lightness (y-axis); the lower rows (C2-6) correspond to desaturated colours at medium lightness. Light, desaturated rectangles indicate low naming consistency at the category border. Circles illustrate whether and how often a colour chip was chosen as a prototype of a colour term. The rightmost vertical bar shows results for achromatic (greyscale) colour chips. The figure is taken from Witzel & Gegenfurtner (2018), which provides further details.

References

Rogers, M. R., Witzel, C., Rhodes, P., & Franklin, A. (2020). Color constancy and color term knowledge are positively related during early childhood. Journal of Experimental Child Psychology, 196, 104825. https://doi.org/10.1016/j.jecp.2020.104825

Siuda-Krzywicka, K., Boros, M., Bartolomeo, P., & Witzel, C. (2019). The biological bases of colour categorisation: From goldfish to the human brain. Cortex. https://doi.org/10.1016/j.cortex.2019.04.010

Siuda-Krzywicka, K., Witzel, C., Bartolomeo, P., & Cohen, L. (2020). Color Naming and Categorization Depend on Distinct Functional Brain Networks. Cerebral Cortex, 31(2), 1106-1115. https://doi.org/10.1093/cercor/bhaa278

Siuda-Krzywicka, K., Witzel, C., Chabani, E., Taga, M., Coste, C., Cools, N., Ferrieux, S., Cohen, L., Seidel Malkinson, T., & Bartolomeo, P. (2019). Color Categorization Independent of Color Naming. Cell Reports, 28(10), 2471-2479.e2475. https://doi.org/10.1016/j.celrep.2019.08.003

Witzel, C. (2018). Misconceptions About Colour Categories. Review of Philosophy and Psychology, 10, 499–540. https://doi.org/10.1007/s13164-018-0404-5

Witzel, C., Flack, Z., Sanchez-Walker, E., & Franklin, A. (2021). Colour category constancy and the development of colour naming. Vision Research, 187, 41-54. https://doi.org/10.1016/j.visres.2021.05.008

Witzel, C., & Gegenfurtner, K. R. (2018). Color Perception: Objects, Constancy, and Categories. Annual Review of Vision Science, 4(1), 475-499. https://doi.org/10.1146/annurev-vision-091517-034231

Is there hierarchical generalization in action-effect learning?

Lea Eichfelder[1], Volker H. Franz[2], Markus Janczyk[3]

[1]Department of Psychology, Universität Bremen
[2]Department of Computer Science, Universität Tübingen

On the role of categorization in evaluative conditioning: Category abstraction increases the generalization of acquired likes and dislikes.

Kathrin Reichmann, Mandy Hütter, Barbara Kaup, Michael Ramscar
University of Tübingen

Attitudes and preferences play a central role for decision making in social contexts, for example when deciding whom to interact with in a group of novel people. In such situations, previously acquired attitudes may be generalized towards novel individuals. The present work studies the underlying mechanisms of generalization. In particular, we consider category formation throughout learning as one possible pathway of generalizing attitudes and propose categorization as a cognitive mechanism that could promote the generalization of attitudes and preferences.

To put this idea to an empirical test, we employed evaluative conditioning (EC) as an experimental model of attitude acquisition and generalization. The EC paradigm allows investigating the mechanisms underlying a change in evaluations of conditioned stimuli (CSs, e.g., neutral pictures) as a function of previous pairings with likable or dislikable stimuli (e.g., positive or negative pictures; for reviews, see De Houwer et al., 2001; Hofmann et al., 2010). Previous research demonstrated that attitudes acquired via EC generalize to novel stimuli. For example, the conditioning of category exemplars as CSs changed evaluations towards novel category members and the whole stimulus category (Glaser & Kuchenbrandt, 2017; Luck et al., 2020; Olson & Fazio, 2006). However, as of yet it is unknown whether the generalization of evaluations towards novel category exemplars may further increase when participants group CSs into categories during conditioning.

In four experiments, we modified the conditioning procedure in a way that one condition facilitated the grouping of CSs into categories that were predictive of positive or negative valence. More specifically, we manipulated the number of CSs of a specific category that were included in the learning phase in a between-participants design. Participants either saw a single CS or multiple different CSs of a category together with stimuli of positive or negative valence. Experiencing multiple different CSs per category should facilitate participants forming CS categories, and attaching evaluative meaning directly to the category-level. Consequently, wider generalization towards novel category exemplars was expected in the former relative to the latter condition.

As predicted, our experiments demonstrated wider generalization towards novel stimuli of a category when multiple rather than a single CS per category were presented during the conditioning procedure. Moreover, recognition memory performance and evaluations of distinct CS parts confirmed the interpretation that the different learning procedures lead to evaluative representations of different levels of abstraction. We discuss practical implications for generalization phenomena in social contexts such as stereotypes and prejudices, and propose predictive learning as a functional mechanism underlying category abstraction in EC.

Keywords: evaluative conditioning, generalization, categorization, predictive learning

References

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative Learning of Likes and Dislikes: A Review of 25 Years of Research on Human Evaluative Conditioning. *Psychological Bulletin, 127*(6), 853–869. https://doi.org/10.1037/0033-2909.127.6.853

Glaser, T., & Kuchenbrandt, D. (2017). Generalization effects in evaluative conditioning: Evidence for attitude transfer effects from single exemplars to social categories. *Frontiers in Psychology, 8*(103), 1–16. https://doi.org/10.3389/fpsyg.2017.00103

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative Conditioning in Humans: A Meta-Analysis. *Psychological Bulletin, 136*(3), 390–421. https://doi.org/10.1037/a0018916

Luck, C. C., Patterson, R. R., & Lipp, O. V. (2020). Be careful what you say!–Evaluative change based on instructional learning generalizes to other similar stimuli and to the wider category. *Cognition and Emotion*, 1–16. https://doi.org/10.1080/02699931.2020.1816912

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32*(4), 421–433. https://doi.org/10.1177/0146167205284004

Towards unifying category learning and probability learning using the CAL framework of rule extrapolation and contextual modulation

Rene Schlegelmilch[1], Andy J. Wills[2], & Bettina von Helversen[1]
[1] **University of Bremen (GER),** [2] **University of Plymouth (UK)**

Abstract

Stimulus classification is an everyday feat (e.g., in medical diagnoses by differentiating ultrasound images). Category feedback, however, is often non-deterministic (e.g., by 25% chance untrue a.k.a. probabilistic feedback) rendering experiences as somewhat unreliable, and the question is how humans (still) learn stimulus-category regularities. In probability learning and economic decisions, however, the question usually reverses to why humans do not perfectly exploit regularities when correct categorization leads to reward (e.g., non-rational probability matching; Feher da Silva, et al., 2017; Plonsky, Teodorescu, & Erev, 2015). Here, we address both questions in a domain-general framework formalizing how humans, in probabilistic tasks, simultaneously learn category regularities and sequential feedback regularities. We use our recently introduced Category Abstraction Learning (CAL) framework (Schlegelmilch, Wills, & von Helversen, 2021), implementing the idea that participants count the streak of common events to predict *when* rare events or violations of a learned regularity will occur (i.e., expectations of non-random probability). CAL not only predicts probability matching in general, but also the proportion of strategies often discussed as Win-Stay-Lose-Shift (WSLS), probability maximizing and more recently studied sequential pattern learning (akin to gamblers fallacy). CAL also provides an account of expectancy priors (see Koehler & James, 2014), proposing that they stem from an awareness that unobserved stimuli lead to unobserved outcomes (contrasting). CAL predicts the same type of behavior in probabilistic fear conditioning, and provides access to the so-called Perruchet effect (Perruchet, 2015), shown to dissociate associative processes (conditioned responding) from more deliberate decision processes (outcome expectancy ratings). One central CAL hypothesis is, that visual and sequential features are processed as outcome predictors (cue-outcome associations) or contextual modulators (conditionals for cue-outcome associations). To evaluate the CAL predictions, we present two reanalyses of a recent probability learning study (Feher da Silva, et al., 2017; see Figure 1) and a fear conditioning study (Lee, Hayes, & Lovibond, 2018). We discuss CAL's central learning hypotheses from both perspectives, arguing that different learning strategies vary on a continuum of CAL's parameters, often leading to conditional hypotheses about non-stationary (context-dependent) outcome probabilities.

Figure 1 shows the results of a reanalysis of a typical probability learning study (Feher da Silva, et al., 2017), in which participants saw 300 trials, in each deciding whether a left or a right box will contain a ball, receiving reward if correctly predicted. The ball appeared in the left box in 70% of the trials (common event; truly random), otherwise in the right box (rare event; mutually exclusive). To evaluate CAL's predictions (shown in the talk), we applied a Bayesian model on the data, splitting the 300 trials into 37 bins, in each assessing the likelihood of five different characteristic behavioral patterns for each participant in a model selection (strategies: FBA - feedback-based adjustment a.k.a. WSLS, maximizing [PC], Guessing, and two sequential rules). The sequential rules (see also Szollosi, Donkin, & Newell, 2022) reflect that expectancy of rare events increases with increasing common-streak length (TPM2), and that rare events occur at a regular rate (TPM1; e.g., three trials after the last rare event), each predicted by CAL.

Keywords: Category Learning, Probability Learning, Rule Abstraction, Feedback Patterns

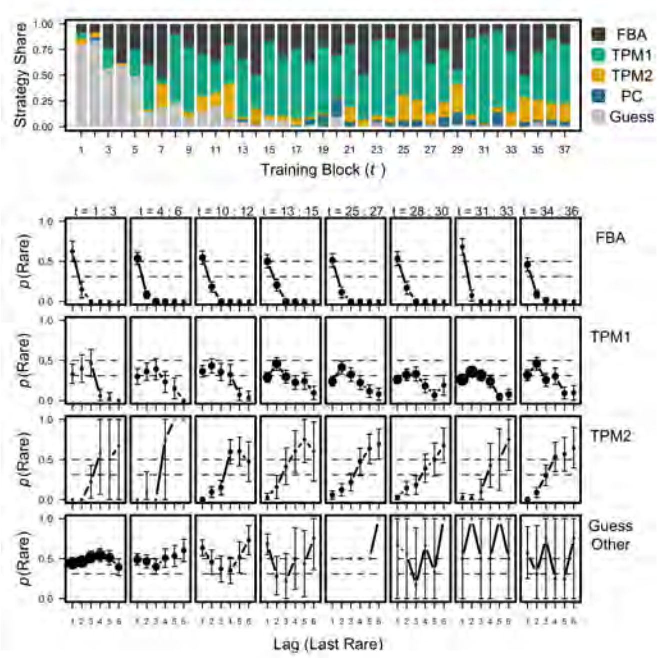


Figure 1: Reanalysis of the probability learning task of Feher da Silva, et al. (2017; see abstract). (Top) Cumulative proportions (sum to 1; y-axis) of participants (N=80) assigned to five strategies in 37 bins (t, eight trials each; x-axis). FBA = feedback-based adjustment (choosing recently reinforced option), TPM1 = temporal probability matching (e.g., choosing rare three trials after observing rare), TPM2 = rare expectancy increasing with distance to last rare event, PC = predict common (maximizing), Guess = 50% responding. (Bottom) Corresponding average choices (rare expectancy by recency [Lag] of last rare event) in selected training blocks (PC is not depicted, as the averages were simply near zero). Symbol size indicates number of observations in each group and at each lag. Error bars = between 95% CI's.

References

(1) Feher da Silva, C., Victorino, C. G., Caticha, N., & Baldo, M. V. C. (2017). Exploration and recency as the main proximate causes of probability matching: a reinforcement learning analysis. *Scientific Reports*, 7 , 15326. doi: 10.1038/s41598-017-15587-z 1

(2) Koehler, D. J., & James, G. (2014). Probability matching, fast and slow. In *Psychology of Learning and Motivation* (Vol. 61, pp. 103–131). Academic Press.

(3) Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018). Peak shift and rules in human generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(12), 1955–1970. doi: 10.1037/xlm0000558

(4) Perruchet, P. (2015). Dissociating conscious expectancies from automatic link formation in associative learning: A review on the so-called perruchet effect. *Journal of Experimental Psychology: Animal Learning and Cognition, 41* (2), 105–127. doi: 10.1037/xan0000060

(5) Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological Review, 122* (4), 621–647. doi: 10.1037/a0039413

(6) Schlegelmilch, R., Wills, A. J., & von Helversen, B. (2021). A cognitive category-learning model of rule abstraction, attention learning, and contextual modulation. *Psychological Review, Advance online publication*. doi:10.1037/rev0000321

(7) Szollosi, A., Donkin, C., & Newell, B. R. (2022). Toward nonprobabilistic explanations of learning and decision-making. *Psychological Review, Advance online publication*. doi: 10.1037/rev0000355

An Embodied Neural Process Model Grounds Structured Representations

Daniel Sabinasz, Gregor Schöner
Institute for Neural Computation at Ruhr-University Bochum

Higher cognitive competences like language comprehension and thought are often assumed to depend on "structured representations", i.e., representations that consist of parts which are combined in a specific way (e.g., Fodor & Pylyshyn, 1988; Jackendoff, 2002). On the other hand, it is generally recognized that higher cognitive competences supervene on "grounded" perceptual and motor representations and associated brain processes (e.g., Barsalou, 2008; Lakoff & Johnson, 1999) that are tied to the sensing and acting body in an environmental context.

Higher cognitive competences then cannot be explained as mere manipulations of symbolic representations that have no association with their grounded meaning; nor can they be explained by embodied dynamical systems or connectionist models as long as these models do not explain the capacity to productively combine representations (Barsalou, 1999; Fodor & Pylyshyn, 1988).

Consider sentences with a subject, a transitive verb, and an object, where the subject and object are described as noun phrases with nested prepositional phrases. Example: "the ball moves towards the big tree, which is to the left of the house and to the right of the lake". We comprehend such sentences by associating the individual words with concepts and combining those concepts in accord with how the words are arranged in the sentence as per the grammatical rules of the language. By these means, we are able to link the sentence to our perceptual experience. This enables us, e.g, to identify the referents of the different noun phrases and identify their relations, which in turn enables us to verify the truth of the sentence.

Following Jackendoff (2002), we postulate that the processing of natural language leads to a neural short-term memory representation of its conceptual structure (Figure 1, left). We propose a novel neural dynamic model of how the brain may represent the conceptual structure of a subject-verb-object sentence (Figure 2), ground its components in prototype concepts (Figure 3), combine those concepts consistent with their arrangement in order to establish reference to objects in a visual scene (not plotted), and effectively verify the truth of the sentence. The model is based on the neural principles formalized in Dynamic Field Theory (Schöner et al., 2015) and is a single dynamical system evolving in continuous time.

For simplified visual scenes and a selected set of concepts, we demonstrate in simulations how the model can verify the truth of arbitrary nested subject-verb-object sentences.

Keywords: concepts; language comprehension; grounding; embodiment; prototype theories; neural process model; conceptual structure; conceptual combination; dynamic field theory

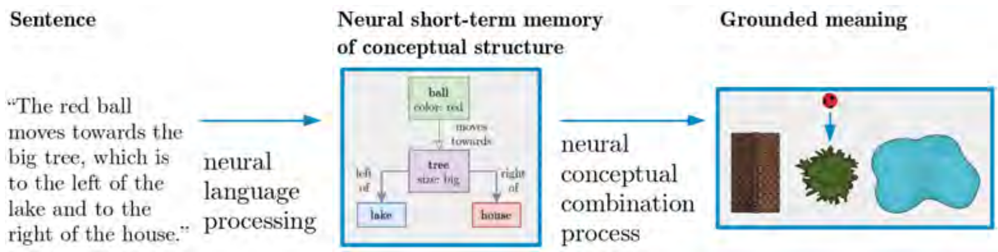


Figure 1: From language through conceptual structure to grounding

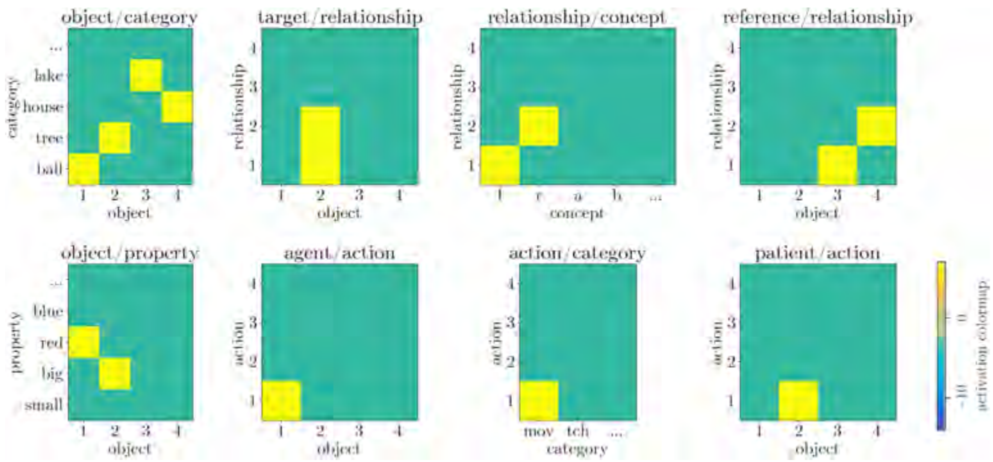


Figure 2: Activation snapshots of conceptual structure short-term memory fields.

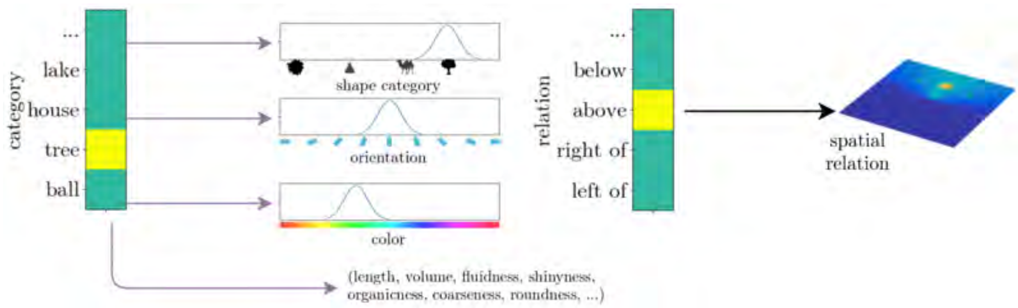


Figure 3: Model of how categorical representations can be grounded in prototype concepts through projections to neural fields defined over visual or conceptual feature spaces

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4).

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1-2).

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought.* Basic Books.

Schöner, G., Spencer, J. P., & the DFT Research Group. (2015). *Dynamic Thinking: A Primer on Dynamic Field Theory*. Oxford University Press.

Symposium 4

Symposium proposal for the 15. Biannual Conference of the German Society for Cognitive Science, Freiburg i.B.

Digital childhood: social robots and social interactions

Adriana Hanulíková (University of Freiburg)

Katharina Rohling (University of Paderborn)

Language acquisition is an important endeavor of childhood: language enables us to communicate with and learn from others. Early language development seems critical for later academic success, and modern society often requires children to learn more than one language. There is solid evidence that the acquisition of linguistic skills is strongly modulated by cognitive and perceptual factors as well as human-to-human interactions. Yet, over the last years, social robots as catalysts for learning start to attract scholarly attention. A future without robots is unthinkable, but the social, educational, and ethical challenges they pose are far from understood.

An important challenge for social robotics is to determine whether and under what conditions robots can help cognitive development and (language) learning during childhood on one hand, and how robots should be designed to achieve an efficient interaction with children on the other hand. By its very nature, social robotics research is fragmented across multiple disciplines (psychology, neuroscience, cognitive sciences, robotics, artificial intelligence, linguistics, among others), making it important to develop a systematic and truly interdisciplinary approach. Understanding how children perceive and interact with robots, how robots should be designed for an efficient use, and how robots affect child development allows the research to arrive at an integrated understanding of social robots in digital childhood.

The aim of this workshop is twofold: First, to provide an overview of social robotic research and explore the options and limits of social robots to foster language and cognitive development in one-to-one interactions and classroom settings; and second, to identify future challenges of social robots for child interactions and learning.

The workshop further aims to stimulate an interdisciplinary discussion by inviting speakers from different academic fields. The talks will address questions such as what is social about learning, how social experience and expectations shape humans' comprehension of and interaction with robots, and how social robots can help to foster learning and development.

The symposium will be opened by an introduction from the organizers and closed by a structured discussion.

Keywords: child development, social robots, learning, digital media, language acquisition

Talk, Listen & Keep Me Company: A Mixed Methods Analysis of Children's Perspectives towards Robot Reading Companions

Nathan Caruana[1], Ryssa Moffat[1], Aitor Miguel-Blanco[1] & Emily S. Cross[1, 2, 3,4]

[1] School of Psychological Science, Macquarie University, Australia
[2] Centre for Elite Performance, Expertise and Training, Macquarie University, Australia
[3] Institute of Neuroscience and Psychology, University of Glasgow, United Kingdom
[4] MARCS Institute for Brain, Behaviour and Development, University of Western Sydney, Australia

The deployment of assistive robots in education is being rapidly realised, with many different types of robots already used to support various types of learning, with children who have different needs and whilst serving different social roles (e.g., robot tutor, peer or novice; Belpaeme et al., 2018). However, the emerging research in this space has neglected to fundamentally examine the features that will make robots appropriate and successful, or to seek the input from children themselves. The current study set out to do this by exploring children's first impressions, expectations and experiences (positive and negative) when reading with different types of social robots. We propose that social robots have the potential to promote reading engagement in children by providing socioemotional support and a non-intimidating, interactive context for learning.

We recruited a sample of 30 children (M_{age} = 8.51 years, *SD* = 1.74; 11 females; 67% Caucasian, 33% Asian) using a mixed methods approach to determine the robot features children feel make a good reading companion. This included quantified measures of robot perceptions and preferences upon viewing three robots which differed in both form and function (NAO, Cozmo, MiRo). All responses were collected using a bespoke web-based application run on an iPad, which enabled children to rate (Figure 1A) and rank (Figure 1B) the robots using an intuitive visual response format.

Children provided a number of ratings, including how intelligent and friendly they found each robot to be, with anthropomorphic form (NAO) and emotive responsivity (Cozmo) likely driving stronger ratings of intelligence and friendliness respectively for these robots (see Figure 2). Children were then asked to select one robot to read with. The robot responded to the child's reading at several pre-determined plot points using a range of appropriate emotive animations that were executed using a Wizard-of-Oz approach. An in-depth interview was conducted following the reading task including additional ratings of the robot's 'performance' (see Figure 3) as well as open-ended questions. Interviews were transcribed and then analysed using an inductive thematic analysis approach (Braun & Clarke, 2019; (see Figure 4 for a summary).

Children explained that robots can provide them with an engaging and non-judgmental social context that promotes reading engagement. This was supported by children's perceptions of robots as being intelligent enough to read, listen and comprehend a story, particularly when the robot could talk or display appropriate emotional reactions. However, a key challenge was optimally timing robot behaviour, which in some cases had the potential to distract children. This is difficult to resolve using either human operators or autonomous algorithms.

We demonstrate that much is to be gained from commencing the design and application of social robotics in education contexts by asking children to share their perspectives. This offers the opportunity to identify the robot features that are likely to help and hinder children when learning alongside artificial companions.

*Keywords***:** reading, human—robot interaction, technology assisted education, agent perception

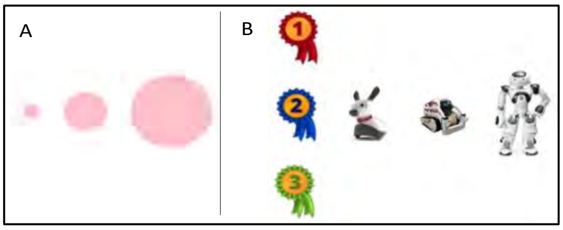


Figure 1. (A) Visual rating scale. (B) Robot preference ranking display screen.

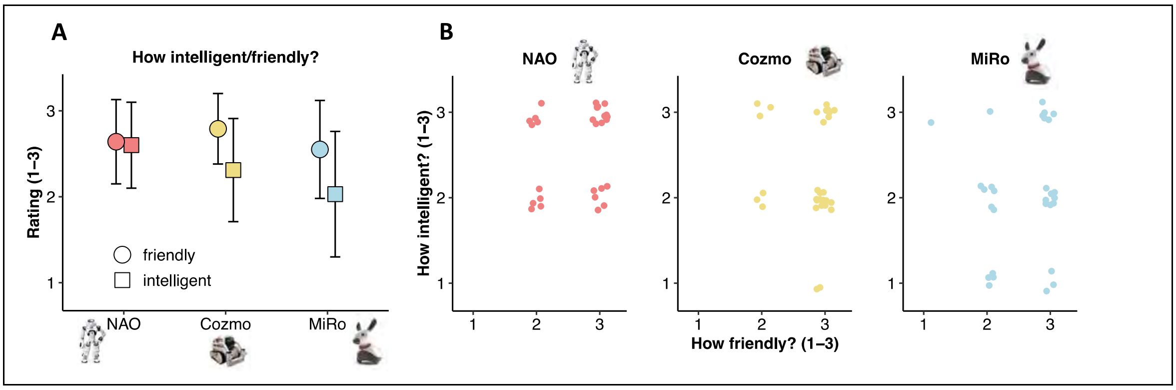


Figure 2. Mean ratings for how ‘intelligent’ and ‘friendly’ children found each robot, plotted separately (A) and against each other (B).

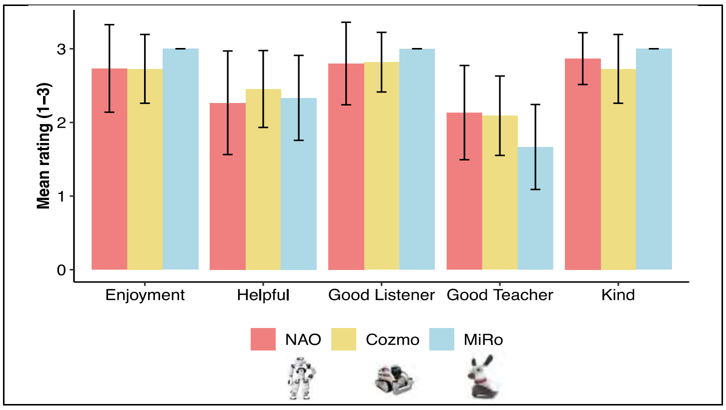


Figure 3. Subjective ratings of robots following reading activity.

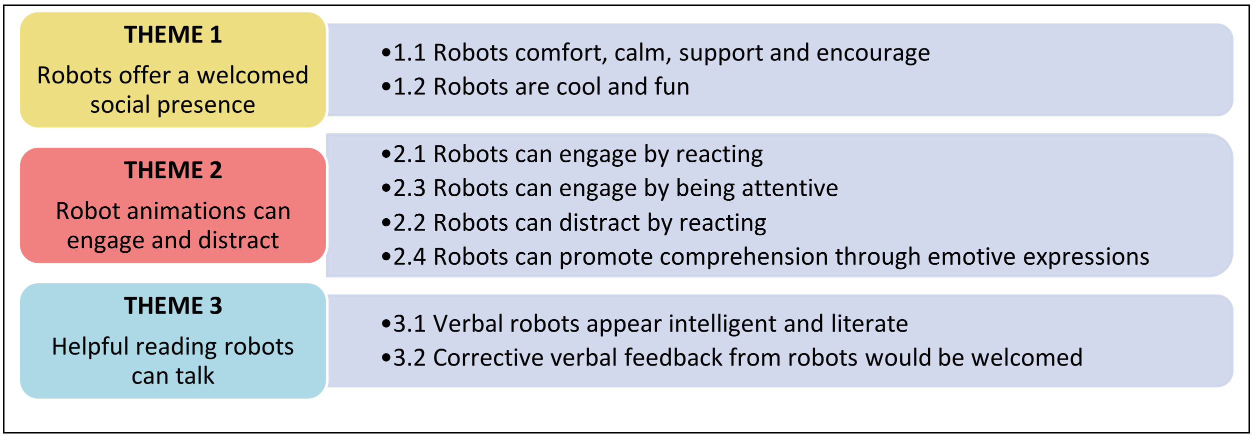


Figure 4. Summary of themes identified through thematic analysis of interviews with children.

References

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*. https://doi.org/10.1126/scirobotics.aat5954

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

“It’s not the robot who learns, it’s me!” considerations on the role of social robots in learning

Barbara Bruno
EPFL, Lausanne

The words 'collaborative learning' describe a situation in which particular forms of interaction among people are expected to occur, which would trigger learning mechanisms, but there is no guarantee that the expected interactions will actually occur. Hence, a general concern is to develop ways to increase the probability that some types of interaction occur. (Dillenbourg, 1999, p. 5). Some of the successful “ways” developed by researchers to increase that probability rely on technology: first in the form of computers (Dillenbourg, 2009) and, more recently, also in the forms of robots (Plauska, 2014) and social robots (Alves-Oliveira, 2014).

Social robots are, arguably, the closest a machine can get to humans: their embodiment is often designed to resemble the human body (Pandey, 2018), their behaviour patterns and even learning algorithms outwardly mimic ours (Penaloza, 2012) and their interaction modalities are carefully crafted to be intuitively familiar, especially concerning emotions expressivity (Tielman, 2014). As such, their use in collaborative learning settings opens previously unconceivable avenues for research, including the idea of people learning by collaborating “with” social robots (Diyas, 2016), rather than “through” them.

Taking the “robot-as-a-peer” paradigm to an extreme (Diyas, 2016), in the first part of my talk I will discuss rationale, methods and curious findings of the case in which a social robot supports a child with handwriting difficulties by motivating him to be the teacher that the robot looks up to, to learn how to write (Hood, 2015).

By observing the interactions between the child and the robot in that scenario, and their evolution over time, we can clearly see their variety and subtlety, as well as their impact on the triggering of desirable learning mechanisms. As an example, let us consider the amount of attention the child gives to the robot: if it’s too little the benefits of the robot’s presence and the collaboration with it are likely to be lost, but if it’s too much the child might end up neglecting the learning activity and thus similarly live a sub-optimal learning experience.

In the second part of my talk I will relate ongoing efforts in investigating how robots’ behaviours can be designed to favour the emergence of collaborative patterns conducive to learning.

*Keywords***:** robots for education, social robots, HRI, collaborative learning.

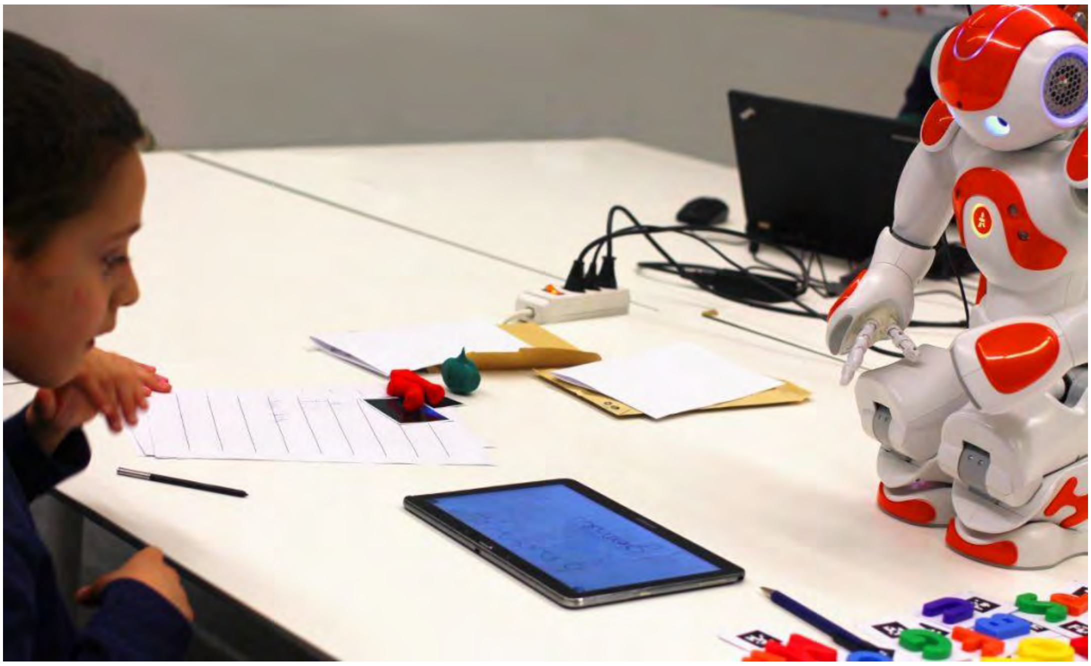


Figure 1. “Learning-by-teaching” with a social robot: a child teaches a robot how to write.

References

Alves-Oliveira, P., Janarthanam, S., Candeias, A., Deshmukh, A., Ribeiro, T., Hastie, H., ... & Aylett, R. (2014, August). Towards dialogue dimensions for a robotic tutor in collaborative learning scenarios. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 862-867). IEEE.

Dillenbourg, P. (1999). What do you mean by collaborative learning?.

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In *Technology-enhanced learning* (pp. 3-19). Springer, Dordrecht.

Diyas, Y., Brakk, D., Aimambetov, Y., & Sandygulova, A. (2016, March). Evaluating peer versus teacher robot within educational scenario of programming learning. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 425-426). IEEE.

Hood, D., Lemaignan, S., & Dillenbourg, P. (2015, March). The cowriter project: Teaching a robot how to write. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts* (pp. 269-269).

Pandey, A. K., & Gelin, R. (2018). A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robotics & Automation Magazine*, *25*(3), 40-48.

Penaloza, C. I., Mae, Y., Ohara, K., & Arai, T. (2012, August). Social human behavior modeling for robot imitation learning. In *2012 IEEE International Conference on Mechatronics and Automation* (pp. 457-462). IEEE.

Plauska, I., & Damaševičius, R. (2014, October). Educational robots for Internet-of-Things supported collaborative learning. In *International Conference on Information and Software Technologies* (pp. 346-358). Springer, Cham.

Tielman, M., Neerincx, M., Meyer, J. J., & Looije, R. (2014, March). Adaptive emotional expression in robot-child interaction. In *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 407-414). IEEE.

Can a social robot advance children's long-term word learning of morphologically complex words by systematically varying the interaction?

Nils Tolksdorf & Katharina J. Rohlfing

Paderborn University

Symposium 5

KogWiss 2020 symposium proposal

Curiosity as a tool for information acquisition

Organizer: Prof. Gert Westermann, Department of Psychology, Lancaster University, UK, g.westermann@lancaster.ac.uk

Summary:

Curiosity - the intrinsically motivated pursuit of knowledge for its own sake - has recently gained substantial interest in Psychology and Cognitive Science more broadly. This is because investigating intrinsically motivated exploratory behaviour shifts focus from the learner as a passive recipient of information to a more autonomous role in which the learner maintains agency and structures the environment in ways that may be conducive to learning. Consequently, a better understanding of how curiosity arises and supports learning has found particular interest in developmental and cognitive psychology, but also in the design of artificial learning systems.

This symposium brings together four experts to present recent research on curiosity from an interdisciplinary perspective. Each talk will last 25 minutes including questions, and the talks will be followed by a general discussion.

In the first talk, Matthias Gruber will discuss the brain bases of curiosity and how states of curiosity enhance learning of both high-curiosity information but also of incidental information presented during states of curiosity. He will argue that the effects of curiosity on memory are mediated by individual differences in the structural-functional connections between dopaminergic regions and the hippocampus.

In the second talk, Gert Westermann will present an artificial neural network model of curiosity-based exploration. The model replicates looking time data from studies with infants and suggests the maximization of predicted learning as a mechanism underlying intrinsically motivated exploration. An analysis of the exploratory sequences generated by the model reveals an overall intermediate complexity, confirming a 'Glodilocks' effect in what information is preferentially learned.

In the third talk, Marina Bazhydai will focus on curiosity-based learning in a social environment. She argues that when faced with epistemic uncertainty, such as when insufficient or unreliable information is provided, preverbal infants query their social partners through social referencing (looking at others to obtain a response) to selectively seek information from more knowledgeable adults to satisfy their curiosity.

In the fourth talk, Lena Ackermann turns to the role of curiosity in word learning as the foundation for language development. While historically, individual differences in toddlers' vocabulary size and structure have been explained in terms of the input a child receives, she describes a series of experiments showing that children's curiosity and interest on the object and category level are strong predictors of novel word recognition and retention in children aged 24 to 36 months, highlighting the importance of curiosity in shaping individual vocabulary trajectories.

Presenters and working titles of presentations:

Matthias Gruber (School of Psychology Cardiff University, UK)
Structural-functional brain connections that support curiosity-related memory enhancements

Gert Westermann (Dept. of Psychology, Lancaster University, UK): *A computational model of curiosity-based learning in infancy*

Marina Bazhydai (Dept. of Psychology, Lancaster University, UK): *Curiosity in social learning: infants' active information seeking from others in epistemic uncertainty*

Lena Ackermann (Psychology of Language, University of Göttingen, DE): *How curiosity shapes early vocabularies*

How curiosity affects learning and information seeking via the dopaminergic circuit

Matthias J. Gruber, CUBRIC, Cardiff University, UK.
Kathrin C.J. Eschmann, CUBRIC, Cardiff University, UK.
Yana Fandakova, Max Planck Institute for Human Development, Berlin, Germany.
Duarte F.M.M. Pereira, CUBRIC, Cardiff University, UK.

Abstract

Over the last decade, research on curiosity – the desire to seek new information – has been rapidly growing. Several studies have shown that curiosity elicits activity within the dopaminergic circuit and thereby enhances hippocampus-dependent learning (e.g., Kang et al., 2009; Gruber et al., 2014). However, given this new field of research, we do not have a good understanding yet of (i) how curiosity-based learning changes across the lifespan, (ii) why some people show better learning improvements due to curiosity than others, and (iii) whether lab-based research on curiosity translates to how curiosity affects information seeking in real life. In this talk, I will present a series of behavioral and neuroimaging studies that address these three questions about curiosity. The first two studies use a trivia paradigm in which participants rate their curiosity about answers to trivia questions (pre-answer curiosity) along with the actual interestingness about the answers (post-answer interest). The level of curiosity-based learning is measured via a surprise memory test for all presented trivia answers. Experiment 1 using the trivia paradigm reveals findings on how pre-answer curiosity and post-answer interest about answers to trivia questions affect learning differently in childhood and adolescence (Fandakova & Gruber, 2021). Findings from Experiment 2 demonstrate how inter-individual differences in the magnitude of curiosity-based learning within the trivia paradigm (i.e., answers to high- compared to low-curiosity questions) depend on the strength of resting-state functional connectivity within the cortico-mesolimbic dopaminergic circuit. Finally, the results of Experiment 3 show how the level of resting-state functional connectivity within this circuit is also associated with the frequency of real-life information seeking (i.e., about Covid-19-related news) (Eschmann, Pereira, et al., in press). Together, our findings help to refine our recently proposed framework – the Prediction, Appraisal, Curiosity, and Exploration (PACE) framework – that attempts to integrate theoretical ideas on the neurocognitive mechanisms of how curiosity is elicited, and how curiosity enhances learning and information seeking (Gruber & Ranganath, 2019). Furthermore, our findings highlight the importance of curiosity research to better understand how curiosity can be harnessed to improve learning and information seeking in real life.

References

Eschmann, K. C. J., Pereira, D. F. M. M., Valji, A., Dehmelt, V., & Gruber, M. J. (in press). Curiosity and mesolimbic functional connectivity drive information seeking in real life. *Social Cognitive and Affective Neuroscience.* (preprint: https://doi.org/10.1101/2022.01.28.478038)

Fandakova, Y., & Gruber, M. J. (2021). States of curiosity and interest enhance memory differently in adolescents and in children. *Developmental Science*, *24*(1), e13005.

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron*, *84*(2), 486–496.

Gruber, M. J., & Ranganath, C. (2019). How Curiosity Enhances Hippocampus-Dependent Memory: The Prediction, Appraisal, Curiosity, and Exploration (PACE) Framework. *Trends in Cognitive Sciences*, *23*(12), 1014–1025.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963–973.

A computational model of curiosity-based learning in infancy

Gert Westermann[1], Katherine Twomey[2]
[1]Lancaster University, UK [2]University of Manchester, UK
g.westermann@lancaster.ac.uk
katherine.twomey@manchester.ac.uk

Curiosity is often conceptualized as driven by metacognitive awareness of a knowledge gap and the drive to close this gap (e.g., Loewenstein, 1994; Pekrun, 2019). This adult-centred perspective has led to claims that infants (and animals), lacking such metacognitive awareness, are not curious (Inan, 2012). Nevertheless, in recent years research on infant curiosity and information seeking has become a focus of early developmental research (for an overview see Bazhydai, Twomey & Westermann, 2020).

Here we present an artificial neural network model to provide a mechanistic account of infant curiosity. The model is based on the principle that infants' active exploration is driven by the maximization of in-the-moment learning progress, which has recently been shown to account for infants' active exploration strategies (Poli et al, 2020). The model was an auto-encoder with an input, a hidden and an output layer that was trained with the backpropagation algorithm and learned to reproduce its input on the output layer. Such networks have previously been used to model infant category learning, and the model's error (difference between input and output) has been used as a proxy for infants' looking times in such tasks (Mareschal, French & Quinn, 2000; Westermann & Mareschal, 2012).

We used a stimulus set consisting of novel animal drawings where exemplars differed across four features (neck and leg length, ear distance, and tail thickness; Fig 1). In a first simulation we modeled a novelty preference study with these stimuli in which 10-month-olds (n = 24) learned categories from maximally, but not minimally, complex stimulus sequences, i.e., sequences in which successive stimuli were maximally dissimilar (Mather & Plunkett, 2011). In a second, crucial simulation we implemented a 'curiosity' mechanism in which the model autonomously selected the next exemplar to learn (instead of the 'experimenter' choosing the next stimulus) on the basis on which of the remaining stimuli afforded maximal learning (weight adaptation).

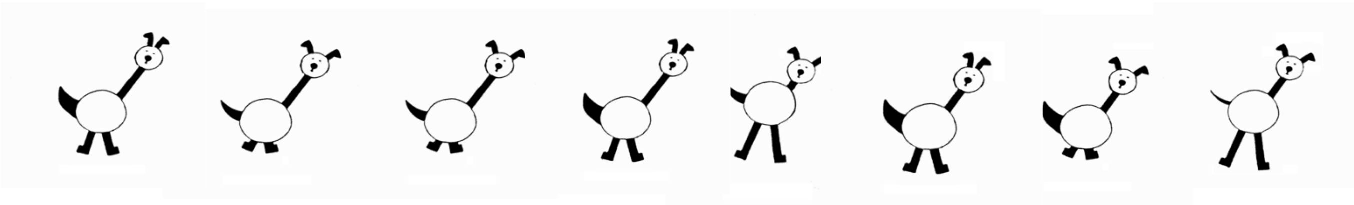
With proportion test error as a proxy for looking time, our first simulation showed strongest categorization after maximally complex sequences (M = 0.99, p < .0001), capturing the empirical data of Mather & Plunkett (2011). However, in the second simulation where the 'curious' model chose its own inputs the network learned equally well (M = 0.97, p < .0001, maximum vs. curiosity: p = 0.28), suggesting that exploration and learning based on learning progress maximization is a powerful mechanism of effective information acquisition.

We analyzed the sequences in which the 'curious' model chose exemplars for learning and found that overall, it selected sequences of intermediate novelty, exhibiting a Goldilocks effect like that found in infant exploration (Kidd, Piantadosi & Aslin, 2012). The model further showed that more similar and more novel stimuli alternated in the exploration sequence.

These simulations represent the first computational investigation of curiosity-based learning in infants, suggesting that in-the-moment learning progress maximization affords optimal learning akin to an adult providing an optimally structured environment. The model makes the novel prediction that infants actively exploring their environment will select stimulus sequences of overall intermediate complexity while varying item-to-item similarity. Overall, this work illustrates active development as a system in which learning is driven by dynamic interactions between the learner and the environment, with the precise nature of these interactions dependent on the learner's current state on the developmental trajectory.

Keywords: curiosity, infants, cognitive development, artificial neural networks, categorization

Figure 1
Stimuli used in the model and experiments



References

Bazhydai, M., Twomey, K., & Westermann, G. (2020). Curiosity and exploration. In Benson, J.B. (Ed.), *Encyclopedia of Infant and Early Childhood Development (2nd ed.), Vol. 2: Cognition, Perception & Language* (pp. 370–378). Academic Press.

Inan, I. (2012). *The Philosophy of Curiosity*. New York, London: Routledge

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLoS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

Mareschal, D., French, R. M., & Quinn, P. C. (2000). A connectionist account of asymmetric category learning in early infancy. *Developmental Psychology*, *36*(5), 635–645. https://doi.org/10.1037/0012-1649.36.5.635

Mather, E., & Plunkett, K. (2011). Same items, different order: Effects of temporal variability on infant categorization. *Cognition*, *119*(3), 438–447. https://doi.org/10.1016/j.cognition.2011.02.008

Pekrun, R. (2019). The murky distinction between curiosity and interest: state of the art and future prospects. *Educational Psychology Review*, *31*(4), 905–914. https://doi.org/10.1007/s10648-019-09512-1

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, *6*(39), eabb5053. https://doi.org/10.1126/sciadv.abb5053

Westermann, G., & Mareschal, D. (2012). Mechanisms of developmental change in infant categorization. *Cognitive Development*, *27*(4), 367–382. https://doi.org/10.1016/j.cogdev.2012.08.004

Curiosity in social learning: Infants' active information seeking from others in epistemic uncertainty

Authors: Marina Bazhydai[1], Gert Westermann[1] and Eugenio Parise[2]
[1]Lancaster University, Lancaster, UK
[2] University of Trento, Italy

Infants are curious learners, actively probing both non-social and social environment for information. As they navigate the world, infants may encounter epistemic uncertainty, such as when insufficient information is provided or prior knowledge is not supported by others' testimony. In response to epistemic uncertainty, preverbal infants' active communication has been proposed to serve an information-seeking function. Among such communicative cues are pointing, social referencing, and babbling. While pointing has been the focus of extensive experimental investigation, less is known about this cognitive function in pre-verbal and pre-pointing infants' social referencing behaviour. Pre-verbal infants' active communicative responses to epistemic uncertainty help delineate core cognitive competencies underlying epistemic development. In two studies, we show that eleven- and twelve-month-old infants initiate active communicative responses to epistemic violation of expectation events and selectively seek unavailable information from more knowledgeable social partners in situations of referential uncertainty. These studies support the proposition that infants in their first year show sensitivity to knowledge distribution among social partners, exhibit epistemic vigilance in social contexts, and actively communicate with interlocutors when they need epistemic input.

In Study 1, 11-month-olds (N = 48) were presented with familiar objects, with caregivers enabling three between-subject conditions: Congruent (caregiver providing a matching label to the object), Incongruent (mismatching label), and No Label. Infants engaged in more social referencing in Incongruent compared to both No label and Congruent condition. These results suggest that epistemic uncertainty upon detecting object-referent mismatch elicited social referencing in pre-verbal infants, indicating an epistemic function of these communicative behaviours. Such enhanced social referencing was present when caregivers were consistent in their incongruent labeling but not when they provided a mix of incongruent and congruent labels, further suggesting that infants formed a coherent representation of their caregiver as a knowledgeable interlocutor and expected reliable and congruent information from them.

In Study 2, we further probed whether infants' social referencing will be selective towards a previously knowledgeable social partner upon facing a different kind of epistemic uncertainty - referential. Twelve-month-olds (N=30) were introduced to two adults, an Informant (reliably labeling objects) and a Non-Informant (equally socially engaging, but ignorant about object labels). Upon being familiarised with the epistemic status of the available interlocutors, infants were asked to locate a novel referent among two novel objects – to make an essentially impossible referential choice ("Which one is the Modi?"). In this situation, infants selectively referenced the Informant rather than the Non-Informant, but showed no such selectivity at other phases of the procedure when no uncertainty was present. These results suggest that infants appropriately formed respective epistemic profiles of available social partners and chose to solicit information from a more knowledgeable one as a better source of information, but only when their input was in fact needed.

Overall, these studies demonstrate that pre-verbal infants are sensitive to epistemic uncertainty and actively communicate to obtain information. Specifically, they generate social looks prior to onset of language or mastery of pointing gesture as part of their developing interrogative communicative toolkit. By actively and selectively seeking information from more knowledgeable others, infants emerge as active participants in social knowledge exchange process.

Keywords: information-seeking, social referencing, infancy, active learning, social learning

How curiosity shapes early vocabularies

Lena Ackermann[1,2,3]
[1] University of Göttingen, Germany, [2] Leibniz ScienceCampus Primate Cognition, Göttingen, Germany,
[3] Radboud University Nijmegen, Netherlands

Why can some children name all animals at the zoo while others know their backhoes from their bulldozers? In this talk, we propose that curiosity is a driving factor in word learning that can explain individual differences in the composition of early vocabularies.

Young children are remarkable word learners: At the end of the first year of life, typically developing children only produce a handful of words, but their vocabularies grow many times over during the second and third year of life. While the overall pattern of vocabulary growth is relatively stable across children and languages, we observe considerable variability with regard to the individual words known to children (Frank et al., 2021). Historically, this variability has been explained in terms of differences in the quantity and quality of input that children receive. Recently renewed interest in the child as an active learner provides a promising backdrop against which we will examine variability in the early lexicon: Children structure their own learning environment by preferentially attending to and learning from stimuli they are curious about (Begus & Southgate, 2012; Begus et al., 2014).

Across three eye-tracking studies, we investigate the influence of interest and curiosity on early word learning and vocabulary composition in children aged two to three years. Curiosity was measured using pupillometry and parental reports. The first study examines whether being curious about a novel object and the semantic category to which it belongs helps 30-month-old children form new word-object associations. We found that children more robustly recognized word-object associations from categories they were more curious about. This result points at a key role of curiosity in early word learning.

Building on these findings, the second study examines the role of curiosity in novel word retention. Children aged 24 and 38 months were tested on their recognition of newly-learnt word-object associations immediately after exposure and with delays of five minutes and 24 hours. We found evidence for a beneficial role of category curiosity especially at 24 months, an age group for which previous evidence of word retention in laboratory word learning tasks was limited (Horst & Samuelson, 2008; Bion et al., 2013).

In the third study, we test whether curiosity guides 30-to-36-month-old children’s referent assignment in a referentially ambiguous word learning situation. Here, we found that referent assignment was guided by relative curiosity about one object over the other, while there was no evidence for a role of category-level curiosity in resolving referential ambiguity.

Taken together, the results of all three studies point at a vital role of interest and curiosity in early word learning: Referent selection, initial word-object mappings, and longer-term word retention are all positively influenced by curiosity at the object or category level. These findings help explain variability in early vocabularies, while also adding to our growing understanding of the importance of curiosity as a tool for information acquisition in early learning and development.

Keywords: curiosity, word learning, language development, individual differences

References

Begus, K., Gliga, T., & Southgate, V. (2014). Infants learn what they want to learn: Responding to infant pointing leads to superior learning. *PLoS ONE*, 9(10), e108817. doi: 10.1371/journal.pone.0108817

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science*, 15(5), 611–617. doi: 10.1111/j.1467-7687.2012.01160.x

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126(1), 39–53. doi: 10.1016/j.cognition.2012.08.008

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). Variability and Consistency in Early Language Learning: The Wordbank Project. MIT Press.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157. doi: 10.1080/15250000701795598

Symposium 6

Title:

Cognition-centered Human-Robot Interaction

Organizers:

Prof. Dr. Nele Rußwinkel, Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Kognitive Modellierung in dynamischen Mensch-Maschine-Systemen

Prof. Dr. Eva Wiese, Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Kognitionspsychologie und Kognitive Ergonomie

Linda Heimisch, M.Sc., Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Kognitive Modellierung in dynamischen Mensch-Maschine-Systemen

Participants:

Prof. Dr. Eva Wiese, Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Kognitionspsychologie und Kognitive Ergonomie:
Working title: Robots as social agents: Challenges and insights from social neuroscience

Linda Heimisch, M.Sc., Technische Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Kognitive Modellierung in dynamischen Mensch-Maschine-Systemen
Working title: What can cognitive modelling do for Human-Robot Interaction?

Prof. Dr.-Ing. Stefan Kopp, Universität Bielefeld, AG Kognitive Systeme und soziale Interaktion
Working title: Models of dynamic social cognition and behavior for human-aware HRI

Dr. Chenxu Hao, Friedrich-Alexander-Universität Erlangen-Nürnberg, Department Elektrotechnik-Elektronik-Informationstechnik, Lehrstuhl für Autonome Systeme und Mechatronik:
Working title: Integrating models of cognitive and physical Human-Robot Interaction

Abstract:

With progressing development in the field of autonomous agents in the context of both social and industrial robotics, the interaction of humans and robots becomes an increasingly important field of research. Psychological research on human-robot interaction in large part focuses on cooperation, collaboration and co-existence of humans and robots. While in this field a substantial ground of research already exists on ergonomics and usability, our symposium aims to focus on a cognitive science perspective on the topic. From a cognitive point of view, the interaction of humans and robots can be thought of as simultaneous

information processing, with one human and one artificial agent involved at the same time. This brings up various essential questions referring to a variety of topics in cognitive science, including knowledge, (semantic) representations, learning, and goal-oriented behavior: Is there an advantage of providing a robot with human-like cognitive abilities, involving human shortcomings, over an AI approach that aims at optimized task solving? If so, to what extent are shared representations of the task environment necessary? Which participant should have what kind of knowledge for a successful division of work? Of particular importance for a trouble-free interaction is that the robot has a solid representation of the goal that its human team mate follows, or a mental model of it, and that it is able to adapt to dynamic situations. From a cognitive perspective, it is therefore essential that the robot is able to identify the intention that the human follows, and changes of intention during task performance. In our symposium we want to address both the potential of a cognitive science perspective for advancing common task solving of humans and robots, as well as the related challenges. The motivation of our symposium is to bring together researchers from different backgrounds. We invite speakers to contribute relevant research approaches from disciplines related to the cognitive sciences.

Robots as Social Agents: Challenges and Insights from Social Neuroscience

Yasmina Giebeler & Eva Wiese
Berlin Institute of Technology

ABSTRACT
Social robots as future cohabitants become more and more reality, and at present they are already used as social companions for elderly people or as therapeutical intervention aids for children with autism spectrum disorder or patients with sensorimotor impairments. Outside the clinical context, social robots foster collaboration in the workplace, or teach science in the classroom. They also facilitate activities in daily lives as friendly assistants in supermarkets and airports. While lots of progress has been made in terms of the technical realization of social robots, their ability to interact with humans in a truly social way is still quite limited. The biggest challenge is to determine how to design social robots that can adapt to the users' cognitive and technical abilities but are also perceived as social companions that understand the needs, feelings, and intentions of their human partners. One way to achieve this goal is using a systematic experimental approach that combines behavioral and physiological neuroscientific methods such as eye tracking, electroencephalography (EEG) or even functional near-infrared spectroscopy (fNIRS) with realistic interaction scenarios involving physically embodied robots. This approach requires understanding of how humans interact with each other, how they perform tasks together and how they develop feelings of social connection over time, and to use these insights to formulate design principles that make robots attuned to the workings of the human brain. An approach like this adds to the current literature, where subjective ratings are the main tool to assess a robot's performance and socialness, as well as user satisfaction. Although subjective measures like rating scales are suitable to capture the quality of a given robot design, they are neither able to predict performance in human-robot interaction nor to inform roboticists how to improve given designs in order to attune them to the human cognitive system.

Together with other neuroroboticists, I put forward the argument that the likelihood of artificial agents being perceived as social companions can be increased by designing them as intentional agents that activate areas in the human brain involved in social-cognitive processing. I will discuss how attributing mental states to robots can positively affect human-robot interaction by fostering feelings of social connection, empathy and prosociality, and how neuroscientific methods can be used to identify design features that can trigger such attributions. I will also present a series of neuroscientific experiments showing that mind attribution to robots positively affects performance in human-robot interaction by enhancing social-cognitive processes like joint attention. Lastly, I will describe circumstances under which attribution of intentionality to robots might be disadvantageous, and discuss challenges associated with designing social robots that are inspired by neuroscientific principles.

Bad robots? Humans rapidly attribute mental states during the perception of robot faces

Martin Maier [1,2], Alexander Leonhardt [1,3], and Rasha Abdel Rahman [1,2,3]
[1] Department of Psychology, Humboldt-Universität zu Berlin
[2] Science of Intelligence, Research Cluster of Excellence, Technische Universität Berlin
[3] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin

The demand for social robots—embodied artificial systems that interact with humans in their daily lives—is expected to increase in the coming years. As different uses for robots are developed (e.g., care, entertainment, customer interaction) many people may soon engage with robots in their private and professional lives. Awareness about which dimensions in a robot support the trust and cooperation of human partners and which create distrust and anxiety may allow roboticists to build robots that are more readily accepted.

Social robots represent an intriguing puzzle: people's explicit opinion is usually that they are mechanical artifacts, yet they are often prepared to interact with them as if they were human, naturally engaging with the characters that robots represent (Clark & Fischer, 2022). Here we investigated how these competing intuitions play out during the perception of robot faces: to what extent does processing of robot faces in the brain reveal attribution of intentionality? We tested this with a phenomenon known from the perception of human faces: reading expressions into objectively neutral faces according to prior knowledge associated with the persons (Abdel Rahman, 2011; Baum et al., 2020; Maier et al., 2022; Suess et al., 2015).

In two preregistered, fully counterbalanced experiments, participants learned positive, negative or neutral back-stories characterizing the behavior of 36 existing humanoid robots (e.g., "reads to the elderly" vs. "interrogates dissidents" vs. "works in a warehouse"). All presented robots had medium human likeness (i.e. with distinct heads and faces with eyes, yet android robots were avoided). In Experiment 1, we collected 60 participants' ratings of facial expression valence before and after learning the information, as well as each robot's trustworthiness after learning. As predicted, ratings of trustworthiness and facial expressions aligned with the associated prior knowledge, with effects more pronounced for negative information. This provides initial evidence that people may "see" good or bad intentions in robot faces.

In Experiment 2, we recorded EEG while 30 participants rated robots' facial expressions. Using linear mixed models, we analyzed single-trial amplitudes of three event-related potential components with known relevance to face perception and processing of emotionally charged stimuli: the N170, associated with early visual perception of faces, the early posterior negativity (EPN), reflecting fast emotional responses to visual stimuli, and the late positive potential (LPP), an index of more elaborate stimulus evaluation. The valence of prior knowledge—especially negative knowledge—influenced amplitudes of the N170 and the LPP, but not the EPN. These results demonstrate that 1) affective knowledge influences processing of robot faces at an early perceptual stage, meaning that people literally *see* intentions in neutral robot faces; and 2) reflexive emotional responses to robot faces are not influenced in the way that responses to human faces would be (Abdel Rahman, 2011; Baum et al., 2020; Schindler & Bublatzky, 2020).

The results show that, on the one hand, people rapidly and automatically attribute mental states to humanoid robots during perception. On the other hand, people are less emotionally affected by robots' potentially good or bad intentions than they would be when interacting with other humans. Processing in the brain thus reflects people's competing intuitions about social robots at different processing stages. These results lend support to the idea that we may not engage with social robots as fully-fledged intentional social agents, but rather as *depictions* of these. The extent to which these effects extend to robots with more strongly varying degrees of human likeness remains to be explored. Our findings contribute to a better understanding of human-robot social interaction and the implications of our relationships with AI for our individual and social futures.

Keywords: social robots, intentionality attribution, top-down effects on perception, EEG

References

Abdel Rahman, R. (2011). Facing good and evil: Early brain signatures of affective biographical knowledge in face recognition. *Emotion, 11*(6), 1397–1405. https://doi.org/10.1037/a0024717

Baum, J., Rabovsky, M., Rose, S. B., & Abdel Rahman, R. (2020). Clear judgments based on unclear evidence: Person evaluation is strongly influenced by untrustworthy gossip. *Emotion, 20*(2), 248–260. https://doi.org/10.1037/emo0000545

Clark, H. H., & Fischer, K. (2022). Social robots as depictions of social agents. *Behavioral and Brain Sciences*, 1–33. https://doi.org/10.1017/S0140525X22000668

Maier, M., Blume, F., Bideau, P., Hellwich, O., & Abdel Rahman, R. (2022). Knowledge-augmented face perception: Prospects for the Bayesian brain-framework to align AI and human vision. *Consciousness and Cognition, 101*, 103301. https://doi.org/10.1016/j.concog.2022.103301

Schindler, S., & Bublatzky, F. (2020). Attention and emotion: An integrative review of emotional face processing as a function of attention. Cortex, 130, 362–386. https://doi.org/10.1016/j.cortex.2020.06.010

Suess, F., Rabovsky, M., & Abdel Rahman, R. (2015). Perceiving emotions in neutral faces: Expression processing is biased by affective person knowledge. Social Cognitive and *Affective Neuroscience, 10*(4), 531–536. https://doi.org/10.1093/scan/nsu088

Abstract submission

Models of adaptive mentalizing for socially-aware HRI

Stefan Kopp
Research Center „Cognitive Interaction Technology“ (CITEC)
Bielefeld University
skopp@techfak.uni-bielefeld.de

Human-robot interaction has made great advancements in the last decades, including interactive learning, language-based communication, or collaborative physical interaction. At the same time, the field is still facing hard challenges in coping with individual differences of users, with the task- and context-related complexities of cooperative behavior in real-world scenarios, or in maintaining user acceptance over the long term. One development that has started (also in the broader A.I. community)) is to underscore the importance of „human or social awareness“, i.e. the development of robots that can understand other agents more deeply at the level of their mental states, can made themselves understandable to others, or can respond to the needs and expectations of others during collaborative interaction. I will argue that, in addition to the currently predominant data-driven approaches in A.I., an embodied cognitive approach is needed that can account for the embodied-agentive (intentional) as well cognitive-affective nature of agents, as well as for the dynamic and continuous adaptation processes taking place within and between agents that come to be interaction partners. I will present work that aims to increase a robot’s ability for adaptive mentalizing (theory of mind) and for using it to increase robustness, efficiency and acceptability of realtime human-robot interaction. I will show results from work on modeling „satisficing social cognition“ that is sensitive to the requirements, benefits and costs of the cognitive processing of a given interactional situation, on using minimal mentalizing in multimodal dialogue, and on co-constructing mutual understanding through adaptive explanations.

Integrating Models of Cognitive and Physical Human-Robot Interaction

Chenxu Hao[a], Nele Russwinkel[b], Daniel F.B. Haeufle[c,d], and Philipp Beckerle[a]

[a]Friedrich–Alexander Universität Erlangen–Nürnberg, Germany
[b]Kognitive Modellierung in Dynamischen Mensch-Maschine-Systemen, TU-Berlin, Germany
[c]Hertie Institute for Clinical Brain Research and Center for Integrative Neuroscience, University of Tübingen, Germany.
[d]Institute for Modelling and Simulation of Biomechanical Systems, Computational Biophysics and Biorobotics, University of Stuttgart, Germany

Human-robot interactions (HRI) research usually focuses either on the cognitive level or on the physical level. However, characteristics of HRI in real-life scenarios often involve both information exchanges on the cognitive level (Mutlu et al., 2016) and force exchanges on the physical level (Haddadin & Croft, 2016). Specifically, the human agent and the robotic agent need to act adaptively by inferring each other's intentions and predicting each other's actions in the given environment (Beckerle et al., 2018; Schürmann & Beckerle, 2020; Ho & Griffiths, 2022). In such dynamic interactions, the robotic agent also needs to be anticipating and flexible in supporting the human agent. Therefore, it is crucial to create unified models of HRI that can capture the dynamic exchang on both cognitive and physical levels, while taking the interaction environment into account (see Kahl et al., 2022, for an example of such a unified model for a single agent).

The goal of our research is to develop a conceptual framework that provides a possibility to connect models of physical and cognitive HRI with situational awareness. In this framework, both the human agent and the robotic agent are modeled with a high-level cognitive layer and a low-level sensorimotor control layer, or a physical layer. The robotic agent's cognitive and physical layers correspond to those of the human agent (Figure 1), which the robotic agent could use to anticipate the human agent on both cognitive and physical levels.

Inspired by Kahl et al. (2022)'s approach, we propose that one possible implementation of the robotic agent's cognitive layer is an ACT-R cognitive architecture (Anderson, 2009). The robotic agent's cognitive layer keeps track of the task information and goals based on the human agent's needs and the situation. This model can be interfaced with the robotic agent's representation of the human agent's sensorimotor system (e.g., Stollenmaier et al., 2020). Such detailed models of the human agent's movement trajectories and interactive forces may be particularly useful in situations where the robot provides precise physical support to patients with motor control deficits.

To illustrate how our framework can be applied to HRI in real-life settings, we focus our discussion on a specific application scenario where a human agent interacts with an assistive robotic agent. The robotic agent anticipates potential risky situations and provide the human agent with different levels of assistance, ranging from a simple alert through observations of the task situation to full physical support.

Keywords: human-robot interaction, anticipatory thinking, modeling HRI

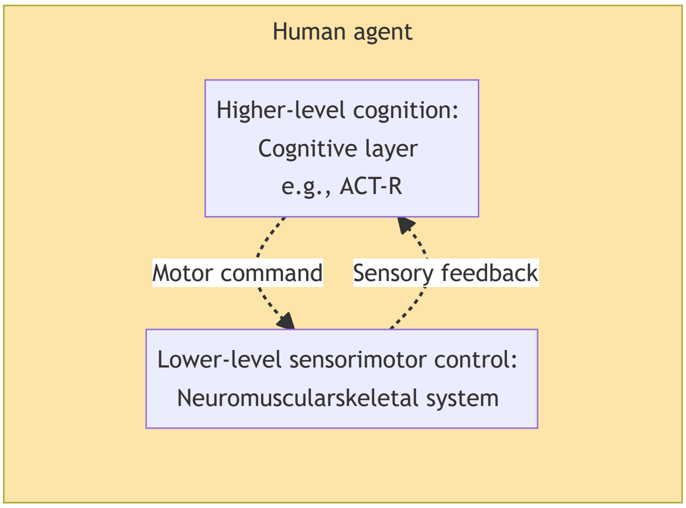


Figure 1: A unified framework connecting the human agent's higher-level cognition and lower-level sensorimotor control. The cognitive layer keeps track of the task information and the situation, which is interfaced with models of the human agent's sensorimotor system on the physical layer. To create unified models of cognitive and physical HRI, the robotic agent may have cognitive and physical layers that correspond to those of the human agent.

References

Anderson, J. R. (2009). *How can the human mind occur in the physical universe?* Oxford University Press.

Beckerle, P., Castellini, C., & Lenggenhager, B. (2018). Robotic interfaces for cognitive psychology and embodiment research: a research roadmap. *Wiley Interdisciplinary Reviews: Cognitive Science*, *10*(2), e1486.

Haddadin, S., & Croft, E. (2016). Physical human–robot interaction. In *Springer handbook of robotics* (pp. 1835–1874). Springer.

Ho, M. K., & Griffiths, T. L. (2022). Cognitive science as a source of forward and inverse models of human decisions for robotics and control. *Annual Review of Control, Robotics, and Autonomous Systems*, *5*(1), 33-53. doi: 10.1146/annurev-control-042920-015547

Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. (2022). Towards autonomous artificial agents with an active self: modeling sense of control in situated action. *Cognitive Systems Research*, *72*, 50–62.

Mutlu, B., Roy, N., & Šabanović, S. (2016). Cognitive human–robot interaction. *Springer handbook of robotics*, 1907–1934.

Schürmann, T., & Beckerle, P. (2020, September). Personalizing Human-Agent Interaction Through Cognitive Models. *Frontiers in Psychology*, *11*, 561510. doi: 10.3389/fpsyg.2020.561510

Stollenmaier, K., Rist, I. S., Izzi, F., & Haeufle, D. F. (2020, November). Simulating the response of a neuro-musculoskeletal model to assistive forces: implications for the design of wearables compensating for motor control deficits. In *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)* (pp. 779–784). New York City, NY, USA: IEEE. doi: 10.1109/BioRob49111.2020.9224411

What can cognitive modelling do for Human-Robot Interaction?

Linda Heimisch, M.Sc.

Technische Universität Berlin

Cognitive modelling is a method in Cognitive Science that allows for the computer-based simulation of human cognitive abilities, not only including different cognitive domains but also ensuring cognitive and empirical plausibility which implies the replication of human errors and delays. Among several established forms and approaches of cognitive modelling, cognitive architectures have emerged as a useful cognitive modelling paradigm particularly in applied human-machine interaction contexts. In my talk I will address the question how cognitive modelling based on cognitive architectures can help to improve collaborative Human-Robot Interaction.

As a formal method that abstracts from whether a concrete cognitive agent is biological or artificial in nature, cognitive modelling is well suited for simulating cognitive processes in contexts where both a human and a robot process information simultaneously to reach a common goal. But who shall be modelled in such environments, and with what benefit? I argue that providing robot agents with cognitive models that mirror a human task-solving approach can bring forward smoothness of collaborative interaction and transparency, including comprehensibility, of a robot's behaviour for its human opposite. Further, I make the claim that for this, cognitive modelling might even be superior to other approaches such as machine learning.

My talk will focus on the cognitive architecture ACT-R which has proven to be a succesful framework for cognitive modelling in Human-Robot Interaction contexts. I will present two studies that are currently being conducted in our lab that use ACT-R models to improve an artifical agent's ability to recognise the intention of a human opposite. I will put a special emphasis on the methodological potentials and limitations of cognitive architectures that we worked out in our projects, with the aim of placing my presentation in the context of a general methodological discussion of different cognitive modeling approaches.

Keywords: Cognitive Modelling, Human-Robot Interaction, ACT-R, Cognitive Robotics.

Symposium 7

Sleeping – theoretical, developmental and applied perspectives on memory, dreams, and creativity

Organizers: Annette Hohenberger and Katharina Lüth, University of Osnabrueck

Sleep plays a major role in our life. When we lack sleep, we are extremely limited in many mental abilities. We can think less logically, make worse decisions, can regulate our emotions less. One explanation for this is our memory restructuring during sleep. Memories are not only replayed and consolidated; they also interconnect in a different way. We can search for new solutions and recognize new patterns during sleep. This happens with everyday tasks and problems we think about, as well as with decisions we make. In this symposium, we want to shed light on the mechanisms during sleep that play a role for memory consolidation in a theoretical, developmental, as well as applied clinical perspective.

Memory replay during sleep and rest: Decoding sequential replay in humans (Simon Kern)

Memory replay and consolidation during wake and sleep has mostly been studied in animals; quantifying replay in humans is still difficult. To tackle this methodological problem, Simon Kern uses a novel brain decoding approach and MEG data to show that previously learned content is rapidly re-visited in a subsequent resting session. This novel research methodology may help further understanding disorders such as Alzheimers, schizophrenia and depression.

The purpose of dreams: an interdisciplinary review (Simone Anthes)

Dreaming is a universal cognitive phenomenon that is difficult to study but interesting to many fields of research. Simone Anthes will provide a review about current and past theories on why we dream and relate them to one another in order to identify similarities and contradictions. She will provide interdisciplinary perspectives and possible points of interest for future studies.

Insightful problem solving in children's tool making: The role of napping (Gökhan Gönül, Anil Karabulut, Annette Hohenberger)

Young children need even more sleep (opportunities) than adults. Here, Gönül, Karabulut and Hohenberger assessed whether napping increases insightful, creative problem solving in preschool children's tool innovation abilities.

Nightmares: Efficacy of the imagery rehearsal therapy *via* mobile phone application (Katharina Lüth)

An important application of theories of sleep and dreams is nightmare therapy. In Imagery Rehearsal Therapy (IRT) against nightmares, creative solution-oriented thinking is trained. Katharina Lüth investigates the effectiveness of IRT *via* a mobile app to facilitate access to nightmare therapy, which is only insufficiently implemented in our health care system. Yet, it is becoming increasingly important, given the mounting psychological stress recently caused by the COVID-19 pandemic and the current war in Ukraine.

Using brain decoding to uncover traces of memory reactivation during rest and sleep

Authors: Kern, Simon (1); Gais, Steffen (2); Feld, Gordon (1,3)
1: Central Institute of Mental Health, Germany;
2: Eberhard Karls Universität Tübingen, Germany;
3: University of Heidelberg, Germany

Replaying memory traces of recently learned material plays an important role for the consolidation of long-term memories (Diekelmann & Born, 2010). The role of replay has been mainly investigated in animal research, whereas measuring replay in humans is still methodologically challenging. In recent years, several approaches using brain decoding methodology have emerged. These approaches leverage pattern similarity of brain activation during exposure to learning material to subsequent patterns appearing in a rest or sleep phase (Belal et al., 2018; Himmer et al., 2021; Schönauer et al., 2017; Schreiner et al., 2018, 2021; Sterpenich et al., 2021). Using different experimental designs, it is possible to further analyze the processes involved in cued and non-cued memory reactivation during rest and sleep. In this talk I will summarize the recent advances and highlight the challenges of these approaches using our own study as an example.

<u>Experiment</u>: In our study 30 participants learned a sequence consisting of ten visually and categorically different visual items while inside an MEG scanner. After recording an eight-minutes resting state as control, we recorded a functional localizer by presenting the ten stimulus items all over again in a pseudo-randomized order for a total of 50 repetitions per item. Then participants learned a (non-semantical, random) sequential ordering of the previously seen stimulus items. After reaching 80% retention performance of the sequence, participants performed an eight-minutes resting state. Finally, participants were tested on their retention performance in a follow-up test.

<u>Analysis:</u> Using the data of the item-presentations during the localizer, we trained machine learning decoders to extract the brain activity patterns that the specific item elicits. Using these decoders, we evaluated the brain activity during the resting state session for evidence of reactivation of individual stimulus items. Using an approach called Temporally Delayed Linear Modelling (Liu et al., 2021), we tested for sequentialness of the decoded items in the same order as the learned sequences, i.e., whether decoded item reactivations follow the same sequential pattern as participants learned during the learning task.

<u>Results:</u> Contrary to our expectation, there was no difference between a control resting state and the experiment resting state in terms of evidence for sequential replay. Using a simulation, we found that physiologically implausible rates of replay (one event per second) would need to be present during the resting state to find a significant effect. That means, our experiment setup was probably not powered enough to detect sparse reactivation, even if it was present. After this insight, we evaluated shorter time spans directly during learning, e.g., after participants acquired new sequence information. Here we find evidence of backwards replay at a time distance of around 100-150 ms between each individual item. However, further analysis is necessary to validate these findings. Last but not least, our results outline the importance of decoder choice and experimental design and the challenges that are posed by using a decoding approach.

Keywords: MEG, sleep, memory replay

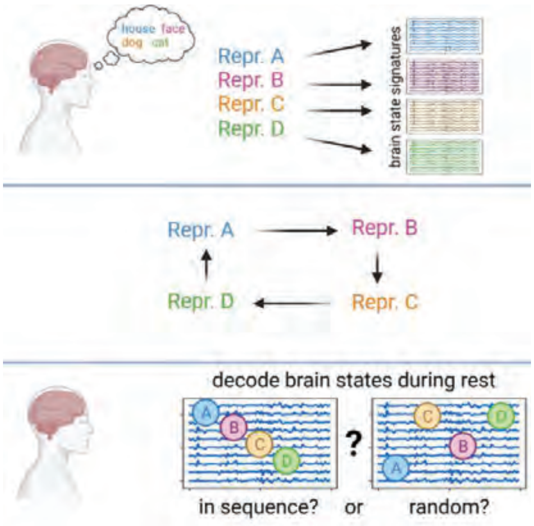


Figure 1: Experimental set-up. In a first step, brain activation patterns of different items are recorded in the MEG. In a second step, participants learn an ordering of the previously seen items. In a third step it is statistically tested whether patterns of step one are re-occuring in the learned order in a subsequent rest or sleep session.

References

Belal, S., Cousins, J., El-deredy, W., Parkes, L., Schneider, J., Tsujimura, H., Zoumpoulaki, A., Perapoch, M., Santamaria, L., & Lewis, P. (2018). Identification of memory reactivation during sleep by EEG classi fi cation. *NeuroImage*, *176*(April), 203–214. https://doi.org/10.1016/j.neuroimage.2018.04.029

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Himmer, L., Bürger, Z., Fresz, L., Maschke, J., Wagner, L., Brodt, S., Braun, C., Schönauer, M., & Gais, S. (2021). *Localizing spontaneous memory reprocessing during human sleep* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.11.29.470230

Liu, Y., Dolan, R. J., Higgins, C., Penagos, H., Woolrich, M. W., Ólafsdóttir, H. F., Barry, C., Kurth-Nelson, Z., & Behrens, T. E. (2021). Temporally delayed linear modelling (TDLM) measures replay in both animals and humans. *ELife*, *10*, e66917. https://doi.org/10/gkf6sk

Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., & Gais, S. (2017). Decoding material-specific memory reprocessing during sleep in humans. *Nature Communications*, *8*(May), 15404. https://doi.org/10.1038/ncomms15404

Schreiner, T., Doeller, C. F., Jensen, O., Rasch, B., & Staudigl, T. (2018). Theta Phase-Coordinated Memory Reactivation Reoccurs in a Slow-Oscillatory Rhythm during NREM Sleep. *Cell Reports*, *25*(2), 296–301. https://doi.org/10.1016/j.celrep.2018.09.037

Schreiner, T., Petzka, M., Staudigl, T., & Staresina, B. P. (2021). Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nature Communications*, *12*(1), 3112. https://doi.org/10/gkgmkx

Sterpenich, V., van Schie, M. K. M., Catsiyannis, M., Ramyead, A., Perrig, S., Yang, H.-D., Van De Ville, D., & Schwartz, S. (2021). Reward biases spontaneous neural reactivation during sleep. *Nature Communications*, *12*(1), 4162. https://doi.org/10/gk5d3h

The Purpose of Dreams

Simone Christiane Anthes
University of Osnabrück

Dreams are phenomena that have been a part of the human experience for millennia and still serve an important role in social and cultural practices (Moss, 2009).
Their position in scientific research, however, is still quite new. In the beginning of the twentieth century, dream analysis was introduced to the western practice of psychoanalysis and psychiatry as a first attempt at the objective study of dreams. However, dream content was very rigidly defined by Sigmund Freud's own interpretation, which stated that dreams represent repressed or unconscious wishes, desires and motivations (Freud, 1900/2014). Finally, with the discovery of Rapid Eye Movement (REM) sleep and the subsequent discovery that eye movements in REM correspond with the dream content, dreams were accepted as verified cognitive phenomena (Dement, 1957). The question remains: what purpose do dreams have? Why did the human mind evolve to hallucinate during nightly unconsciousness and what role do dreams play in the context of other evolutionary processes?
Many theories have been created that attempt to answer these questions. One of the best known theories is an extension of the early psychoanalytic view on dreams. It was theorized that dreams themselves are the language in which our subconscious mind tries to communicate with us and that we would do good in trying to decipher these messages (Jung, 1962). The Continuity Hypothesis sees dreams as a continuation of our daily life: people, environments, and situations that are present in our daily life reoccur in our dreams (Pesant & Zadra, 2006). Another perspective is provided by the simulation theories wherein dreams serve the purpose of simulating either social situations or threats, so that we are prepared when we actually encounter them in the real world (Revonsuo et al., 2015). Others claim that we dream in order to forget. The overfitted brain hypothesis states that bizarre images in dreams prevent the brain from overfitting, which offers a new and interesting perspective for both neuroscientific and machine learning applications (Hoel, 2021). Finally, another question arises: Do dreams even have a specific evolutionary purpose or is it possible that they simply are epiphenomena of processes in the brain that happen during sleep, such as pattern reactivation and memory or emotional consolidation, and have no primary function of their own?

The aim for this presentation is to collect and to present and compare such theories on the evolutionary purpose of dreams in order to identify common aspects and to explore emerging patterns and points of focused interest.

Keywords: dreaming, consciousness, altered states of consciousness, evolution of dreaming

References

Dement, W. & Kleitman, N. (1957). The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *Journal of Experimental Psychology, 53*(5), 339–346. https://doi.org/10.1037/h0048189

Freud, S. (2014). *Die Traumdeutung* in *Sigmund Freud: Gesammelte Werke (pp 7-410)*. Köln: Anaconda. (Original work published 1900)

Hoel E. (2021). The overfitted brain: Dreams evolved to assist generalization. *Patterns (New York, N.Y.)*, *2*(5), 100244. https://doi.org/10.1016/j.patter.2021.100244.

Jung, C. G., & Jaffé, A. (1962). *Erinnerungen, Träume, Gedanken von CG Jung*. Zürich: Rascher

Moss, R. (2009). *The secret history of dreaming*. Novato, CA: New World Library.

Pesant, N. & Zadra, A. (2006), Dream content and psychological well-being: A longitudinal study of the continuity hypothesis. *J. Clin. Psychol.*, 62: 111-121. https://doi.org/10.1002/jclp.20212

Revonsuo, A. & Tuominen, J. & Valli, K. (2015). The Simulation Theories of Dreaming: How to Make Theoretical Progress in Dream Science - A Reply to Martin Dresler. In T. Metzinger & J. M. Windt (Eds). Open MIND: 32(R). Frankfurt am Main: MIND Group. doi: 10.15502/9783958570894.

Insightful problem solving in children's tool making: The role of napping

Gökhan Gönül, Cognitive Science Centre, University of Neuchâtel
Anil Karabulut, IMT School for Advanced Studies Lucca, University of Lucca
Annette Hohenberger, Institute of Cognitive Science, University of Osnabrück

Background: Whereas young children are avid tool users, they experience great difficulties innovating a novel tool to solve an unfamiliar problem (Beck, et al., 2011; Cutting, et al., 2011). Few factors are known to improve preschoolers' tool innovation abilities, among them social learning, hierarchical structuring, and divergent thinking (Gönül, et al., 2018). While a positive effect of napping on young children's memory and executive functions is established (Spencer, 2021), not much is known about the role of sleep in insightful tool innovation. Here, we explore whether napping helps pre-schoolers to gain insight into a tool problem.

Method: 69 3-to-5-year-old preschool children (36 male) attended the study in their kindergartens. They were given the "hook task" (Cutting, et al., 2011) in which they had to assemble a short and a long stick into a hook with which they could reach into a tall jar and lift up a small bucket with a sticker inside (see Fig. 1). Children were assigned to one of two experimental (G1+G2) or two baseline groups (G3+G4), which matched each other for time of Posttest and Pretest, respectively (see Fig. 2). In the first group (G1: Break-Nap) the order of events was: Pretest – Nap – Posttest (Time: 12:30-15:30); (G2: Break-Wake): Pretest – break – Posttest (9:00-12:00); (G3: After nap): Nap – Posttest (13:00-15:30); (G4: Morning): Pretest – Posttest (9:00-9:30). In the Pretest, children received the hook task and were encouraged to use the provided material. Only if they could not solve the task (Phase 1), they were admitted to the further study. In the Posttest, if children still could not solve the task (Phase 2), they were shown the assembled tool (Phase 3, social learning); if they still could not solve the task, they were shown how to manufacture the tool (Phase 4, social learning). Scores of 1-4 were given depending on the phase in which children succeeded (4 for Phase 1; 3 for Phase 2; 2 for Phase 3; and 1 for Phase 4). In addition, as suggested by the literature, gender was considered (Gönül, et al., 2021), and questionnaire data on napping duration as well as night sleep quality and quantity (Horvath, & Plunkett, 2018) was collected.

Results and Discussion: After comparing the two nap *vs* break groups, and also the two break groups, various Generalized Linear Models (GLM, Ordinal Regressions) were applied, with group, gender, and napping duration as the major predictors of tool innovation scores. In accordance with the literature, the rate of spontaneous tool innovation (4 children in Phase 1) was very low and, in Phase 2, only two children from the two break groups spontaneously innovated a hook. Children's performance improved with age, as expected, i.e., older children solved the task in earlier phases of the experiment. Considering the four groups, no difference between them in terms of tool-innovation success was found – probably due to the very low success rate in Phase 2. Yet, napping duration was a successful predictor: Children who habitually napped shorter achieved higher scores as compared to children who napped longer (even after controlling for age) – indicating that not the fact that a child napped but rather how long s/he napped was important. Social learning (in Phases 3 & 4) facilitated success, i.e., children readily used the cues given by the experimenter. Gender also turned out significant: boys achieved reliably higher scores than girls in some of the models.

Conclusion: Insight and innovation in the domain of tools are very rare phenomena that might not be readily observable in a medium-sized sample as ours of habitually napping preschoolers. In addition, 3-5 years of age might be too early to succeed in the hook task – in particular, if the tool needs to be combined from parts. Napping duration as a potential indicator of brain maturity might facilitate performance in a social learning context. Despite inconclusive results, our developmental study provides important hints at calibrating these parameters carefully, and at including observational data such as napping quality and quantity. The role of sex in tool innovation as well as other potential factors need further attention as well.

Keywords: *tool innovation, insight, napping, development, problem solving*

Figure 1. Hook task (left) and tools (right)

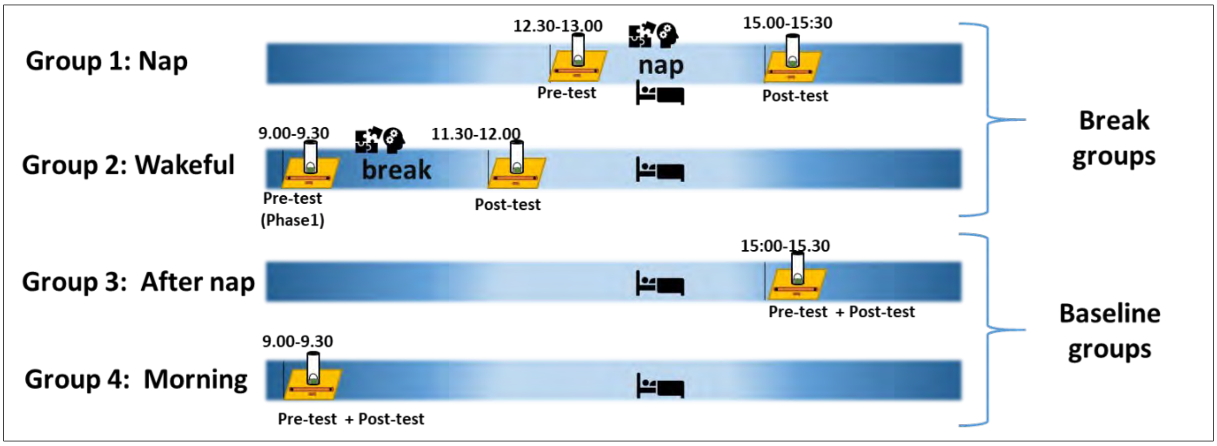


Figure 2. Experimental groups and design

References

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition, 119*(2), 301-306. https://doi.org/10.1016/j.cognition.2011.01.003

Cutting, N., Apperly, I. A., & Beck, S. R. (2011). Why do children lack the flexibility to innovate tools? *Journal of Experimental Child Psychology*, *109*(4), 497-511. https://doi.org/10.1016/j.jecp.2011.02.012

Gönül, G., Takmaz, Kamer, E., Hohenberger, A., & Corballis, M. (2018). The cognitive ontogeny of tool making in children: The role of inhibition and hierarchical structuring. *Journal of Experimental Child Psychology, 173*, 222–238. https://doi.org/10.1016/j.jecp.2018.03.017

Gönül, G., Takmaz, E., & Hohenberger, A. (2021). Preschool children's use of perceptual-motor knowledge and hierarchical representational skills for tool making. *Acta psychologica*, *220*, 103415. https://doi.org/10.1016/j.actpsy.2021.103415

Horváth, K., & Plunkett, K. (2018). Spotlight on daytime napping during early childhood. *Nature and Science of Sleep, 10*, 97-104. https://doi.org/10.2147/NSS.S126252

Spencer, Rebecca M.C. (2021). The role of naps in memory and executive functioning in early childhood. *Advances in Child Development and Behavior, 60*, 139-158. https://doi.org/10.1016/bs.acdb.2020.08.004

A Digital Intervention combining Nightmare Treatment and Insomnia Treatment

Katharina Lüth[1], Gordon Pipa[1], Annika Gieselmann[2]

[1] Institute of Cognitive Science, Osnabrück University, Germany
[2] Department of Clinical Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

Background

Nightmares are a common sleep disorder with an incidence of 5% in the population (Schredl, 2010). Several studies show that nightmares are undertreated and underdiagnosed: People suffering from nightmares rarely seek and receive help (Nadorff et al., 2015; Schredl, 2013), although there are methods such as Imagery Rehearsal Therapy (IRT), whose high efficacy has been shown in meta-analyses (Augedal et al., 2013; Hansen et al., 2013). In IRT, creative solution-oriented thinking is trained. This cognitive treatment is very easy to conduct. Therefore it is an urgent goal to widen its use in our health care system.

Nightmares can occur alone, but are often comorbid with psychological or sleep-related disorders, such as insomnia, which is one of the most common sleep disorders. The treatment of choice for this is cognitive behavioural therapy for insomnia (CBTI) (Riemann et al., 2017).

Since many people suffer from both disorders simultaneously, the combination of CBTI and IRT has been tested and shown to be helpful (Margolies et al., 2013; Ulmer et al., 2011). However, these studies did not compare this combination with the individual treatments. In addition, preliminary data from our research group show that CBTI alone - to some degree - is also helpful for nightmares. Thus, it remains an open question whether IRT combined with CBTI is more effective for people suffering from both disorders, than CBTI alone.

Method

In our ongoing project, we investigate the latter question. We will expand an already existing mobile application against insomnia to include a nightmare therapy module that works according to the principle of Imagery Rehearsal Therapy. Subsequently, the app will be tested for its effectiveness. Parameters such as sleep quality, nightmare frequency, nightmare distress, insomnia severity and depressive symptoms will be assessed before and after using the application. We will include patients suffering from insomnia and nightmares and randomly assign them to three groups: (1) CBTI alone, (2) CBTI and IRT, (3) active control group. We hypothesize that the combination of techniques will have a greater effect on nightmare symptoms than insomnia therapy alone. We also hypothesize that both groups will experience a similar decrease in insomnia symptoms.

App-based Therapy

Telemedicine services such as mobile apps or Internet-based programs can be a partial remedy of the problem that sleep disorders are often undertreated. Since this new form of therapy is currently growing strongly and gaining awareness, it is especially important to accompany this development scientifically. However, the limitations of telemedicine and smartphone-based therapy will also be addressed.

*Keywords***:** Nightmares, Insomnia, Digital health applications, IRT, CBT-I

References

Augedal, A. W., Hansen, K. S., Kronhaug, C. R., Harvey, A. G., & Pallesen, S. (2013). Randomized controlled trials of psychological and pharmacological treatments for nightmares: a meta-analysis. *Sleep Medicine Reviews*, *17*(2), 143-152.

Hansen, K., Höfling, V., Kröner-Borowik, T., Stangier, U., & Steil, R. (2013). Efficacy of psychological interventions aiming to reduce chronic nightmares: a meta-analysis. *Clinical psychology review*, *33*(1), 146-155.

Margolies, S. O., Rybarczyk, B., Vrana, S. R., Leszczyszyn, D. J., & Lynch, J. (2013). Efficacy of a cognitive-behavioral treatment for insomnia and nightmares in afghanistan and iraq veterans with PTSD. *Journal of Clinical Psychology*, *69*(10), 1026–1042.

Nadorff, M. R., Nadorff, D. K., & Germain, A. (2015). Nightmares: under-reported, undetected, and therefore untreated. *Journal of Clinical Sleep Medicine*, 11(7), 747-750.

Riemann, D., Baum, E., Cohrs, S., Crönlein, T., Hajak, G., Hertenstein, E., ... & Spiegelhalder, K. (2017). S3-Leitlinie Nicht erholsamer Schlaf/Schlafstörungen. Kapitel „Insomnie bei Erwachsenen“ (AWMF-Registernummer 063-003), Update 2016, *Somnologie*, 21:2–44

Schredl, M. (2010). Nightmare frequency and nightmare topics in a representative German sample. *European archives of psychiatry and clinical neuroscience*, 260(8), 565-570.

Schredl, M. (2013). Seeking professional help for nightmares: A representative study. *European Journal of Psychiatry*, 27(4), 259–264.

Ulmer, C. S., Edinger, J. D., & Calhoun, P. S. (2011). A multi-component cognitive-behavioral intervention for sleep disturbance in Veterans with PTSD: A pilot study. *Journal of Clinical Sleep Medicine*, *7*(1), 57–68.

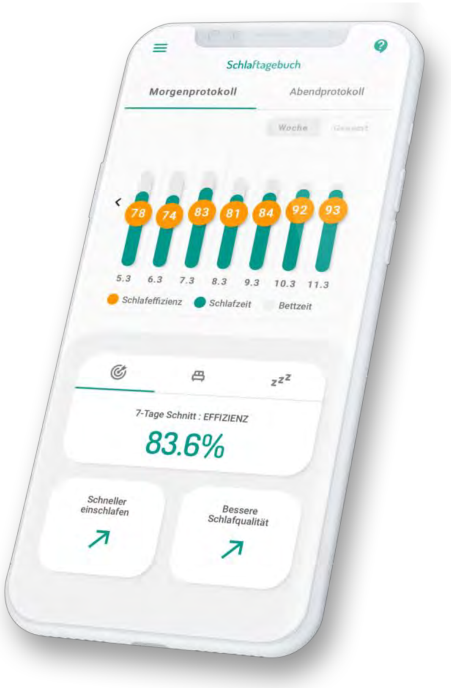


Figure 1: Example of the application somnio. Here, the development of sleep efficiency over several days is displayed, which was created on the basis of sleep diary entries. From: https://somn.io

Sessions

Session 1

In search of lost memories: modeling forgetful generalization

Sebastian Breit[1], Michiko Sakaki[1], Kou Murayama[1], and Charley M. Wu[1]

[1]Eberhard Karls Universität Tübingen

From choosing a restaurant to taking a walk, many situations require us choose between exploring novel options to gain new information or exploiting familiar options known to produce rewarding outcomes. But in situations with a vast number of alternatives, it is not enough to decide between exploration or exploitation. One must also decide where to explore and which experiences are most relevant to inform our decision.

Past work has modeled learning and exploration in large decision-spaces through reward generalization, where past observations can inform novel options on the basis of similarity (Wu et al., 2018). For example, restaurants in similar locations or with similar features (Wu et al., 2020) might produce similar culinary experiences. While this work uses Gaussian Process regression as a Bayesian model of reward generalization, it falls under the general framework of episodic Reinforcement Learning (Gershman & Daw, 2017). These both have in common, that they assume that predictive generalizations about novel options are computed by comparing all past episodes in memory.

Yet, is it really possible to retain every single past experience? Can anyone really remember every single restaurant they have visited? An important cognitive constraint is the limited capacity for memory. In order to make informed choices and to both learn and explore effectively, we need to prioritize which memories to retain.

Here, we introduce a novel method for modeling reward generalization under working memory limitations. To simulate forgetting, we use heteroscedastic Gaussian process regression, where input-dependent noise is added to all observations, representing the decay of memory as a regression to the prior (similar to Collins & Frank, 2012, but with generalization). We model the level of forgetting as a weighted feature vector, where we show that recency- and surprise-based prioritization of memory leads to distinct patterns of choices. Through simulations, we show both model results and parameter estimates are recoverable.

Ongoing work aims to use these models to extend our understanding of the developmental trajectory of learning and exploration (Giron et al., 2022; Mata et al., 2013). In particular, we compare differences between younger and older adults, since previous work has suggested a decline in visuo-spatial working memory around 60 years of age (Ma et al., 2014; Noack et al., 2012). Using a within-subject manipulation of memory load, we will can diagnose how memory limitations contribute to changes in learning and exploration.

Keywords: exploration, generalization, working memory, aging

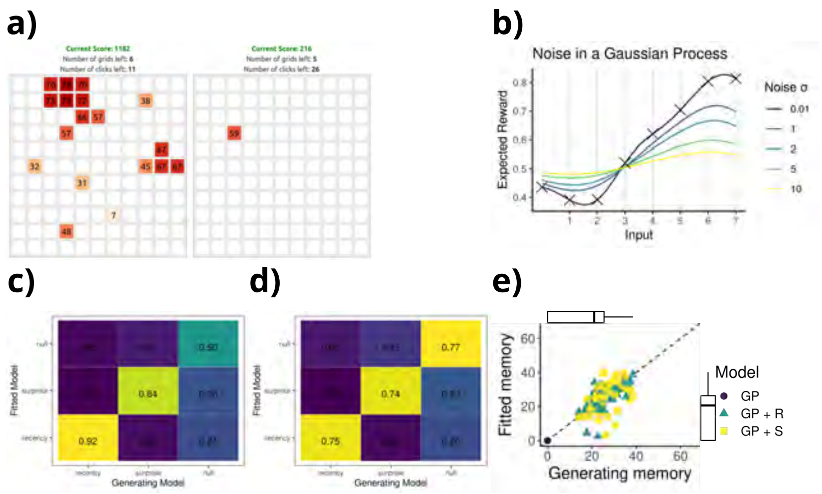


Figure 1: **Experiment Illustration and Model Recovery**. Participants explore a spatially correlated grid trying to gain as many points as possible. Memory load is varied within subjects in each trial. **a)** Low memory load - participants can see all prior choices. High memory load - prior choices disappear after 400ms. **b)** Simulation of the effect of noise (color gradient) on the estimated value (y-axis) of observations (x-axis). **c)** Confusion matrix, with probabilities that a model best predicts the data (y-axis), given an underlying generative model (x-axis). **d)** Inversion matrix, with probabilities that a model did generate the data (x-axis), given that it best fit the data (y-axis). **e)** Correlation between the estimated parameter values (y-axis) and the generating ones (x-axis).

References

Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–1035.

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual review of psychology*, *68*, 101–128.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2022). Developmental changes in learning resemble stochastic optimization. *PsyArXiv*. Retrieved from `psyarxiv.com/9f4k3` doi: 10.31234/osf.io/9f4k3

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, *17*(3), 347–356.

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: is there a reduction in exploration with aging? *Frontiers in neuroscience*, *7*, 53.

Noack, H., Lövdén, M., & Lindenberger, U. (2012). Normal aging increases discriminal dispersion in visuospatial short-term memory. *Psychology and aging*, *27*(3), 627.

Wu, C. M., Schulz, E., Garvert, M. M., Meder, B., & Schuck, N. W. (2020). Similarities and differences in spatial and non-spatial cognitive maps. *PLoS computational biology*, *16*(9), e1008149.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*, 915—924. doi: 10.1038/s41562-018-0467-4

Eliciting everyday beliefs using random generation tasks

Pablo León-Villagrá[1] Lucas Castillo[2] Nick Chater[3] Adam Sanborn[2]

[1]Brown University, USA
[2]University of Warwick, UK
[3]Warwick Business School, UK

Non-symbolical Bayesian reasoning with proportions and probabilities

Katharina Loibl & Timo Leuders
University of Education Freiburg
katharina.loibl@ph-freiburg.de

Bayesian reasoning tasks require the processing of data in probabilistic situations to revise risk estimations (see example task at the top of Figure 1). Research has shown that such tasks are difficult when data is presented in terms of single-event probabilities (Gigerenzer & Hoffrage, 1995; McDowell & Jacobs, 2017). The multiplicative combination of priors and likelihoods can be missed, resulting in erroneous strategies such as prior neglect or averaging heuristics (Cohen & Staub, 2015; Shanteau, 1975). Proportions (i.e., relative frequencies, part-whole ratios) are computationally equivalent to probabilities because they also require the multiplicative combination which is characteristic of normalized data. However, proportions are connected to natural mental representations (so-called ratio sense, Matthews & Ellis, 2018). More specifically, mental representations of nested proportions (e.g., 70% of 20%) allow for a mental operation that corresponds to a multiplicative combination of percentages.

In normalized Bayesian situations, the prior values (i.e., prior probabilities or base rates) and conditional values (i.e., likelihoods or hit rates) have to be combined multiplicatively to derive the posterior values (this does not necessarily mean formal calculation, but also qualitative combination: “a small proportion of a large proportion”). However, such a multiplicative interaction is not intuitive and is therefore cognitively demanding. Unsurprisingly, research shows that humans often fail to apply the Bayes rule correctly but apply biased strategies that deviate from standard Bayesian updating, such as neglecting part of the information, so-called priors-only or evidence-only strategies (Gigerenzer & Hoffrage, 1995). Moreover, when considering all pieces of information, people tend to combine the information additively as opposed to multiplicatively (Shanteau, 1975).

In our study, we systematically varied the conceptual framing (cf. Figure 1, screenshots on l.h.s. and r.h.s.) by adapting the wording across two conditions (e.g., “proportion of…” vs. “probability of …”). We aimed to avoid procedural calculations with percentages and instead focused on the conceptual understanding underlying Bayesian reasoning by utilizing graphical representations without numbers (cf. Leuders & Loibl, 2020). Participants (future teachers) were presented with several situations in which a student has a misconception with varying prior probabilities/base rates. These values were presented again in two separate bars, with the likelihoods/hit rates of an error in a test question given in two adjoining bars. The participants could use movable markers to support their thinking. Finally, they had to position a marker to represent their estimation of a posterior value. In order to classify the participants with respect to their updating strategies, we evaluated their posterior estimations (i.e., their selected positions in the posterior bars) in eight estimation tasks. As participants were not expected to hit the exact posterior position, we followed a naïve Bayesian classification procedure (Duda, Hart, & Stork, 2012; Leuders & Loibl, 2020).

A Bayesian contingency table test (Bayesian equivalent to χ^2-test) revealed moderate evidence (BF_{10} = 6.05) for our hypothesis: participants in the proportion framing condition tended to apply a valid Bayesian strategy (cf. Table 1) while participants in the probability framing condition tended to combine the prior and conditional probabilities additively (by averaging priors and likelihoods as opposed to multiplicatively combining them).

Thus, proportions can be regarded as a natural view on normalized Bayesian situations which may support the learning and understanding of formal Bayesian reasoning.

Keywords: numerical cognition; Bayesian reasoning; biased strategies; proportions; probabilities

Example task (both conditions): A considerable part of students (20%) has a whole-number misconception of decimals (cf. blue part in first bar), i.e., they consider the digits after the decimal point to be whole numbers (WN): These students think that 1.86 > 1.9, because 86 > 9. Of these students, a large part (70%) actually answers this test question erroneously (cf. first grey part). A small part (10%) of students without such a misconception produce an error as well (cf. second grey part). Which proportion (percentage) of students with an erroneous answer to this test question actually has a whole-number misconception (to be indicated in the far right bar)?

condition 1: probability framing

condition 2: proportion framing

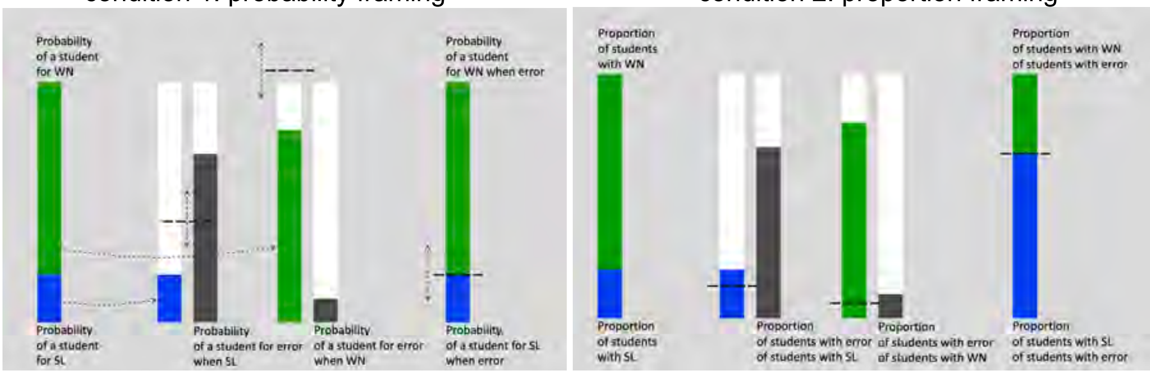


Figure 1. Decision task and graphical environment with probability framing (left) and proportion framing (right).

Table 1. Number of participants per condition classified into the different strategies with high certainty (Bayes factor, $BF_{1:2}$ > 3).

	Bayesian strategy	**Additive strategy**	Priors only	Evidence only
Proportion framing	**13**	**3**	0	2
Probability framing	**5**	**8**	1	0

References

Cohen, A. L., & Staub, A. (2015). Within-subject consistency and between-subject variability in Bayesian reasoning strategies. *Cognitive psychology, 81*, 26-47.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological. Review, 102*, 684–704

Leuders, T., & Loibl, K. (2020). Processing probability information in non-numerical settings – teachers' bayesian and non-bayesian strategies during diagnostic judgment. *Frontiers in Psychology.*

Matthews, P. G., & Ellis, A. B. (2018). Natural alternatives to natural number: The case of ratio. *Journal of numerical cognition, 4*(1), 19-58.

McDowell, M., & Jacobs, P. (2017). Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological Bulletin, 143*(12), 1273-1312.

Shanteau, J. (1975). Averaging versus multiplying combination rules of inference judgement. *Acta Psychologica*, 39, 83-89.

Do verbally processed stimulus-affect contingencies establish stimulus-affect associations?

Torsten Martiny-Huenger
UiT The Arctic University of Norway

Language (e.g., instructions) enables humans to learn and acquire novel responses without direct experiences with the verbally processed information. The mechanisms mediating this ability are debated (e.g., Brass et al., 2017). The present research is concerned with whether verbally processed stimulus-affect contingencies result in associations between the stimulus and the affect. To address this question, I will summarize the results of five experiments illustrating the effects of verbally processed stimulus-affect contingencies on indirectly measured stimulus-induced affect (valence). Regarding the verbal stimulus-affect contingencies, in the experimental condition of all experiments, verbal instructions contained a contingency between a target stimulus (e.g., cupcakes) and an affective concept ('disgusting'). The contingencies were presented to participants as if-then plans (Gollwitzer & Sheeran, 2006) that they repeated to themselves and memorized (e.g., "If I see a cupcake, then I will think disgusting"). The affective consequences were measured in valence-based response-compatibility paradigms (implicit association test, affective priming). The central idea of these paradigms is to measure stimulus-induced valence based on behavioral responses. In both paradigms, participants had to categorize the valence of clearly positive and negative words in the presence of the target or control stimuli. A positive affective response towards the targets is indicated by quicker responses (or fewer errors) to the positive items than to the negative items. In turn, a negative affective response is indicated by quicker responses (or fewer errors) to the negative items than the positive items.

The central hypothesis was that verbally linking a negative affective response ('disgusting') to a target (e.g., cupcakes) leads to a response bias indicating more target-activated negativity (/less positivity) compared to the control conditions/control stimuli.

The presented results are based on convenience samples with 20-50 participants per experimental condition. The hypotheses/procedures of two experiments were pre-registered (Exps. 3 & 4). The hypotheses were tested using an implicit association test (Exps. 1-4) and an affective-priming procedure (Exp. 5) as the valence-based response-compatibility paradigm. The experimental target-disgust condition was either compared with a condition linking a positive affective response to the targets (i.e., 'delicious'; Exps. 1 & 2) or control conditions that presented the targets as something negative (i.e., unhealthy; Exps. 3-5) without an apparent verbal stimulus-affect contingency. Finally, the consequences of the verbal stimulus-affect contingency were tested with a short (<5 minutes) interval between the processing of the verbal information and the valence measure (Exps. 1, 2, & 5) or with a ~24h delay (Exps. 3 & 4).

There is some variability in the results of the tests across the five experiments (based on categorical p-value interpretation). However, overall, the results of all five experiments indicate effects in the hypothesized direction (based on comparing confidence intervals): Participants who memorized the verbal stimulus-disgust contingency showed a response bias that indicated more negativity (/less positivity) towards the targets than towards the control stimuli or participants in the control conditions.

I will discuss the presented results concerning their value in providing evidence that processing verbal stimulus-affect contingencies lead to some form of (procedural) associative learning. With "procedural," I mean that the observed effects are mediated by experience-based ("grounded") representations instead of by some form of language-like, symbolic representations. For example, the effects were observed only when the verbal information contained an apparent stimulus-affect contingency (i.e., target-disgust plan) compared to negativity conveyed without clear "stimulus-response" links. The compatibility effects tended to be stimulus-specific; they did not generalize to responses when the target was not visually present (Exps. 3-5). Finally, the critical verbal information was never directly relevant to executing the response-compatibility paradigms. Thus, the observed consequences of the verbal information were likely to be activated unintentionally. In conclusion, the results can be interpreted as evidence that verbally processed stimulus-affect contingencies can have similar consequences – potentially mediated by similar mechanisms – as directly experienced stimulus-affect contingencies.

Keywords

Stimulus-response learning, Verbal processing, Instruction implementation, Affect, Response-compatibility paradigms

References

Brass, M., Liefooghe, B., Braem, S., & De Houwer, J. (2017). Following new task instructions: Evidence for a dissociation between knowing and doing. Neuroscience & Biobehavioral Reviews, 81, 16–28.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. Advances in Experimental Social Psychology, 38, 69–119.

Session 2

Framing Effects and Numeral Modification

Berry Claus
Leibniz Universität Hannover

Framing variants in risky-choice framing and attribute framing (RCF and AF, see Table 1) studies typically include bare numerals (e.g. 200 people will be saved/400 people will die). The present study investigated whether framing effects (FEs) are modulated by numeral modification. It was conducted in German and involved three numeral modifiers: one that enforces a precise reading (i.e. the German equivalent of *exactly*) and two upper-bound modifiers.

Previous findings on **numeral modification with *exactly*** from RCF studies are mixed. Mandel (2014) found no FE, consistent with his proposal that FEs with bare numerals depend on numeral interpretation and may result from a lower-bound reading (e.g. at least 200 will survive/at least 400 will die), which is blocked when the numerals are modified with *exactly*. However, Chick et al. (2016) observed a robust FE with numerals modified with *exactly*. In Expt. 1 of the present study, effects of numeral modification with *genau* 'exactly' were investigated for RCF and AF. For RCF, it employed a modified German version of a decision scenario adopted from Mandel (2014): war-torn region, in which the lives of 600 people are at stake and in which there are two response plans, a sure option and a risky option (Plan A and B in Table 1, respectively). The outcomes of the response plans were stated in terms of number of people either to be saved (positive frame) or to die (negative frame). For AF, it employed a financial allocation scenario (adopted from Duchon et al., 1989). Participants' task was to indicate whether they would approve or reject the request of an R&D team for additional funding. The frame manipulation pertained to the team's previous performance and was stated in terms of number of either successful (positive frame) or unsuccessful projects (negative frame). For both RCF and AF, there was a significant FE. For RCF, there was a preference for the sure option in the positive-frame condition and a preference for the risky option in the negative-frame condition. For AF, participants more often chose the approval option in the positive-frame condition than in the negative-frame condition. The results of Expt. 1 indicate that enforcing a precise reading of numerals via modification with *genau* 'exactly' does neither preclude AF nor RCF effects. However, they do not exclude that it may reduce FEs.

Expt. 2 (RCF) and Expt. 3 (AF) addressed the effects of two **upper-bound modifiers**, the German equivalents of ***at most* and *up to***. Both modifiers set an upper bound (maximally n). Yet, they exhibit a sharp contrast in evaluative contexts (Blok, 2015; e.g. *Fortunately, at most/up to 100 people will attend my wedding*). Hence, if numeral interpretation is decisive, the two modifiers should not differentially modulate FEs. However, if valence evaluation is the primary source of FEs, then the choice patterns should differ between the two modifiers. The material of Expt. 2 and 3 was the same as in Expt 1, except that the numerals were modified with *höchstens* 'at most' and with *bis zu* 'up to' (see Table 2). In Expt. 2, there was a significant interaction of modifier and frame, with a reversed RCF pattern for *höchstens* (positive frame: preference for risky option; negative frame: preference for sure option) and a standard RCF pattern for *bis zu*. The modifier-by-frame interaction effect was replicated in Expt. 3 for AF.

The findings of this study are challenging for the numeral-interpretation account (Mandel, 2014), but also for other accounts of FEs. The findings could be captured in a dual-process account, when assuming that FEs emerge from the initial valence appraisal of the salient complement of a two-sided issue. By default, the salient complement is that the given predicate (e.g. *saved* or *die*) holds for some instances. However, downward-entailing modifiers, e.g. *at most*, may make the opposite complement salient, i.e. that there are instances for which the predicate does not hold (cf. psycholinguistic accounts of *quantifier focus effects*, e.g. Moxey & Sanford, 2000), possibly resulting in reversed or reduced FEs. All numeral modifiers may potentially affect subsequent deliberate processes, e.g. argumentative and persuasive inferences and quantity implicatures, with varying effects depending on modifier type.

*Keywords***:** framing effects, numeral interpretation, numeral modification, quantifiers

Table 1. Method and results of Expt. 1

Participants	52 native speakers of German	
Material (translated from German)	**RCF**[a]	If Plan A is adopted exactly 200 people will be saved / exactly 400 people will die. If Plan B is adopted, there is a probability of exactly one-third that 600 people will be saved and a probability of exactly two-thirds that nobody will be saved / there is a probability of exactly two-thirds that 600 people will die and a probability of exactly one-third that nobody will die.
	AF[b]	Exactly 30 of the last 50 projects of the team have been successful. Exactly 20 of the last 50 projects of the team have been unsuccessful.
Design	one-factorial within-subject design; counterbalanced orders of RCF and AF blocks and of frame conditions; blocks and conditions were separated by distractor tasks	
Results	**RCF: % of sure option choices** Positive: 52%, Negative: 35%; b=1.43, SE=.65, z=2.22, p<.05	
	AF: % of approvals Positive: 92%, Negative: 65%; b=15.43, SE=4.23, z=3.65, p<.001	

[a]**Risky-choice framing**: Frame manipulation pertains to the wording of two choice options, one with a certain outcome and the other with two outcomes with different probabilities. | [b]**Attribute framing**: Frame manipulation pertains to the description of the attribute of an entity.

Table 2. Method and results of Expt. 2 (RCF) and Expt. 3 (AF)

Participants	101 (Expt. 2) and 94 (Expt. 3) native speakers of German		
Material (translated from German)	**Expt. 2 (RCF)**	If Plan A is adopted at most/up to 200 people will be saved / at most/up to 400 people will die. If Plan B is adopted, there is a probability of maximally one-third that 600 people will be saved and a probability of at least two-thirds that nobody will be saved / there is a probability of maximally two-thirds that 600 people will die and a probability of at least one-third that nobody will die.	
	Expt. 3 (AF)	At most/Up to 30 of the last 50 projects of the team have been successful. At most/Up to 20 of the last 50 projects of the team have been unsuccessful.	
Design	2(MODIFIER) × 2(FRAME) mixed design; random assignment to modifier conditions; counterbalanced order of frame conditions, separated by distractor tasks		
Results	**Expt. 2 (RCF)**	***At most*: % of sure option choices** Positive: 42%, Negative: 65% ***Up to*: % of sure option choices** Positive: 59%, Negative: 45%	Interaction MODIFIER × FRAME b=-1.35, SE=.64, z=-2.09, p<.05
	Expt. 3 (AF)	***At most*: % of approvals** Positive: 67%, Negative: 71% ***Up to*: % of approvals** Positive: 89%, Negative: 69%	Interaction MODIFIER × FRAME b=-1.91, SE=.87, z=-2.19, p<.05

References

Blok, D. (2015). The semantics and pragmatics of directional numeral modifiers. In S. D'Antonio, M. Moroney, & C. R. Little (eds.), *Proceedings of Semantics and Linguistic Theory 25* (pp. 471-490).

Duchon, D., Dunegan, K., & Barton, S.L. (1989). Framing the problem and making decisions: the facts are not enough. *Transactions on Engineering Management, 36*, 25-27.

Mandel, D.R. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General, 143*, 1185-1198.

Moxey, L.M. & Sanford, A.J. (2000). Prior expectation and the interpretation of natural language quantifiers. *Applied Cognitive Psychology*, *14*, 237-255.

How morphological family size affects word recognition in reading and hearing

Hanno M. Müller*,, Louis ten Bosch*, Mirjam Ernestus*,**
*Centre for Language Studies, Radboud University, Erasmusplein 1, 6525 HT Nijmegen, Netherlands
**Department of English and American Studies, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf, Germany

A word's family size (FS) is the type count of all words in which a given word appears as a morphological constituent (Schreuder & Baayen, 1997). For instance, the family of 'play' includes words like 'playful', 'replay', and 'playground'. Words with bigger FSs elicit shorter response times (RT) in visual lexical decision (LD; e.g. Baayen et al., 1997; Kuperman et al., 2009). For auditory LD, reported FS effects are less clear: facilitative (Wurm et al., 2006; Winther Balling & Baayen, 2008), absent (Baayen et al., 2007), or inhibitory (Winther Balling & Baayen, 2012). The latter study finds a facilitative effect for family members that are onset-aligned with the target word indicating that differences between the visual and the auditory FS effect might be due to systematic differences between written and spoken word recognition: Written words likely are perceived at once, whereas spoken words are perceived incrementally during the acoustic signal's unfolding in time.
In this study, because of the conflicting findings on the auditory FS effect, we examined the effect in more detail using different FS definitions and we analyzed the interaction between the words' morphological structures and the FS effect. For analyzing this interaction, a large-scale data set is required, which we have. In addition, we compared visual and auditory word recognition to rule out that differences across the modalities were due to methodological differences. Using generalized additive models (Wood, 2006), we predicted RTs of 20 subjects to 1,932 Dutch unique auditory words from the *Biggest Auditory LD Database Yet* (BALDEY; Ernestus & Cutler, 2015) and of 39 subject to 9,472 unique written words from the *Dutch Lexicon Project* (DLP; Keuleers et al., 2010). Our baseline model included the most prominent predictors except FS that were shown to predict visual and auditory LD RTs in previous studies.
We tested whether including FS (e.g., Schreuder & Baayen, 1997) or onset-aligned FS (Winther Balling & Baayen, 2012) as predictor improved the baseline in terms of AIC. In addition, we defined semantic FS as the type count of family members weighted by their cosine similarity with the target word based on a Dutch word2vec model (Nieuwenhuijse, 2018) and onset-aligned semantic FS as the onset-aligned family members weighted by their cosine similarity. We examined the interaction between the respective predictors and the morphological structure of the words, which were either simplex, consisted of a prefix and a stem, or a stem and a suffix.
The analysis of both the visual and auditory data show that each FS predictor improves the model fit. The two data sets differ in which FS yields the greatest model improvement. For the visual data, semantic FS provides the best fit: the semantic FS has a facilitative effect for all words but the effect size differs per structure. For the auditory data, onset-aligned FS provides the best fit and yields a facilitative effect. However, if the interaction with morphological structure is included in the model, FS and semantic FS return a better fit than onset-aligned FS with a facilitative effect for simplex and suffixed words. Our results indicate that FS not only plays a role in *visual*, but also in *auditory* LD, whereby the effect of a given word's morphological structure varies across the modalities. We think these differences emerge from systematic differences between reading and hearing words.

Keywords: auditory word recognition, visual word recognition, lexical decision, family size, morphological processing

References

Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics*, *35*, 861—877.

Baayen, R. H., Wurm, L. H., & Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon*, *2*(3), 419–463.

Ernestus, M., & Cutler, A. (2015). Baldey: A database of auditory lexical decisions. *Quarterly Journal of Experimental Psychology*, *68*(8), 1469–1488.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 dutch mono-and disyllabic words and nonwords. *Frontiers in psychology*, *1*, 174.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876.

Nieuwenhuijse, A. (2018). *Dutch word2vec model.* Retrieved 2022-02-10, from `https://github.com/coosto/dutch-word-embeddings`

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of memory and language*, *37*(1), 118–139.

Winther Balling, L., & Baayen, R. H. (2008). Morphological effects in auditory word recognition: Evidence from Danish. *Language and Cognitive Processes*, *23*(7-8), 1159–1190.

Winther Balling, L., & Baayen, R. H. (2012). Probability and surprisal in auditory comprehension of morphologically complex words. *Cognition*, *125*(1), 80–106.

Wood, S. N. (2006). *Generalized additive models: An introduction with R*. chapman and hall/CRC.

Wurm, L. H., Ernestus, M. T., Schreuder, R., & Baayen, R. H. (2006). Dynamics of the auditory comprehension of prefixed words: Cohort entropies and conditional root uniqueness points. *The Mental Lexicon*, *1*(1), 125–146.

Animacy Outweighs Topichood when Choosing Referential Expressions

Markus Bader and Yvonne Portele
Goethe University Frankfurt

Languages provide various expressions for referring, including definite descriptions, pronouns and demonstratives. How speakers choose referential expressions is an active area of psycholinguistic research (Vogels et al., 2019). So far, research has focused on the referential expressions used for the sentence subject, ignoring choices for other arguments. That referential expressions are not chosen independently of each other is suggested by the famous Rule 1 of Centering Theory (Grosz et al., 1995, p. 214), rephrased in a non-technical way in (1).

(1) *Rule 1 of Centering Theory*: When any referent is realized by a pronoun, then the topic referent must be realized by a pronoun also.

According to this rule, the sequence *[Peter]$_{TOP}$ called Mary late in the evening. He$_{TOP}$ asked Mary for money.* is pragmatically well-formed whereas the sequence *[Peter]$_{TOP}$ called Mary late in the evening. Peter$_{TOP}$ asked her for money.* is not. Reading time evidence for this contrast comes from Gordon et al. (1993). In a prior picture description experiment with two animate referents, participants adhered to this rule with few exceptions (Portele & Bader, AMLaP 2021 Poster). It is known, however, that animate referents are more often pronominalized than inanimate referents (Fukumura & van Gompel, 2011). We therefore ran an experiment requiring participants to describe pictures with an animate agent and an inanimate patient.

Twenty-four participants read 24 contexts consisting of three sentences, followed by a picture together with a verb (see Table 1). They had to provide a written picture description consisting of a single sentence using the given verb. All verbs required an animate agent and an inanimate patient argument, the most common pairing for transitive events (e.g., Hopper & Thompson, 1980). The preceding context established either the agent or the patient of the depicted event as topic. The third context sentence was either an SO or an OS sentence in order to control for the linear position of the topic. All target sentences produced by the participants were scored according to the referential expressions used for subject and object. The resulting frequencies were statistically analyzed using mixed-effects logistic regression.

Table 2 shows the percentages of referential expressions for the agent (subject) and patient (object) argument of the description. Most references to the agent were made with a pronoun. Pronominal references were much less frequent for patients. For them, about half of all references were made using a definite NP. In addition, more pronouns were used when the referent's antecedent was a subject and a topic. In sum, animacy turned out to be the main determinant of pronominalization, modulated by a small effect of the antecedent's topichood. Demonstratives were used more often for non-topics and for patients/objects than for agents/subjects, in accordance with prior research on demonstratives (e.g., Bosch & Umbach, 2007).

As shown in Table 3, a pronominal subject in combination with a definite object occurred in about 30%, both for agents and patients with topical antecedents. Thus, even when the subject (= agent) was non-topical and the object (= patient) was the topic, participants used pronouns for the subject and definite NPs for the object, contradicting Rule 1 of Centering Theory.

In sum, our results show that animacy as an inherent property of a referent outweighs topichood when it comes to the use of referential expressions, although the latter clearly has its effects. This shows the importance of studying the production of referential expressions in a broader array of structural and semantic configurations.

Keywords: language production, referential expressions, topic structure, picture description

Agent antecedent = topic (topic of final context sentence):
Der Reporter. Bei unserer Zeitung arbeitet ein bekannter Reporter. Dieser Reporter hat schon die sonderbarsten Geschichten erlebt.
SO: Vor kurzem faszinierte ein wertvoller Pokal [den Reporter$_{\text{top}}$].
OS: Vor kurzem faszinierte [den Reporter$_{\text{top}}$] ein wertvoller Pokal.
'*The reporter. A well-known reporter works for our newspaper. This reporter has experienced all kind of strange stories. A short while ago a precious trophy fascinateed [the reporter$_{\text{top}}$].*'

Patient antecedent = topic (topic of final context sentence):
Der Pokal. Unser Verein besitzt einen wertvollen Pokal.
Dieser Pokal stand jahrelang vergessen in einer Rumpelkammer.
SO: Vor kurzem entdeckte zufällig ein bekannter Reporter [den Pokal$_{\text{top}}$].
OS: Vor kurzem entdeckte [den Pokal$_{\text{top}}$] zufällig ein bekannter Reporter.
'*The trophy. Our club owns a precious trophy. This trophy had been standing forgotten in a boxroom for years. A short while ago a well-known reporter discovered [the trophy$_{\text{top}}$] by chance.*'

fotografieren
'to photograph'

Table 1: Sample stimulus. The picture on the right had to be described after reading the context. The agent of the pictured event is printed in blue in the context, the patient in orange.

Referent		Argument with Topic Antecedent: Agent	Patient
Agent	def	15.7	20.2
	dem	1.0	8.9
	pro	83.3	70.9
Patient		Agent	Patient
	def	44.6	54.2
	dem	24.0	6.4
	pro	31.4	39.4

Table 2: Percentages of referential forms used for the agent and patient of the picture description.

Combination	Argument with Topic Antecedent: Agent	Patient
S:def O:def	12.3	16.3
S:dem O:def	0.5	6.9
S:pro O:def	31.9	31.0
S:pro O:dem	22.2	4.9
S:pro O:pro	29.4	35.0

Table 3: Percentages of combinations of subject expression and object expression. Only combinations with a percentage greater than 3% in at least one condition are shown.

References

Bosch, P., & Umbach, C. (2007). Reference determination for demonstrative pronouns. *ZAS Papers in Linguistics*, *48*, 39–51.

Fukumura, K., & van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, *26*(10), 1472–1504.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311–347.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*, 203-225.

Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, *56*, 251-299.

Vogels, J., Krahmer, E., & Maes, A. (2019). Accessibility and reference production: The interplay between linguistic and non-linguistic factors. In J. Gundel & B. Abbott (Eds.), *The Oxford Handbook of Reference* (pp. 337–364). Oxford: Oxford University Press.

Unintentional Response Priming from Verbal Action–Effect Instructions

Yevhen Damanskyy[1] · Torsten Martiny-Huenger[1] · Elizabeth J. Parks-Stamm[2]

1 UiT The Arctic University of Norway, Tromsø, Norway
2 University of Southern Maine, Portland, Maine, USA

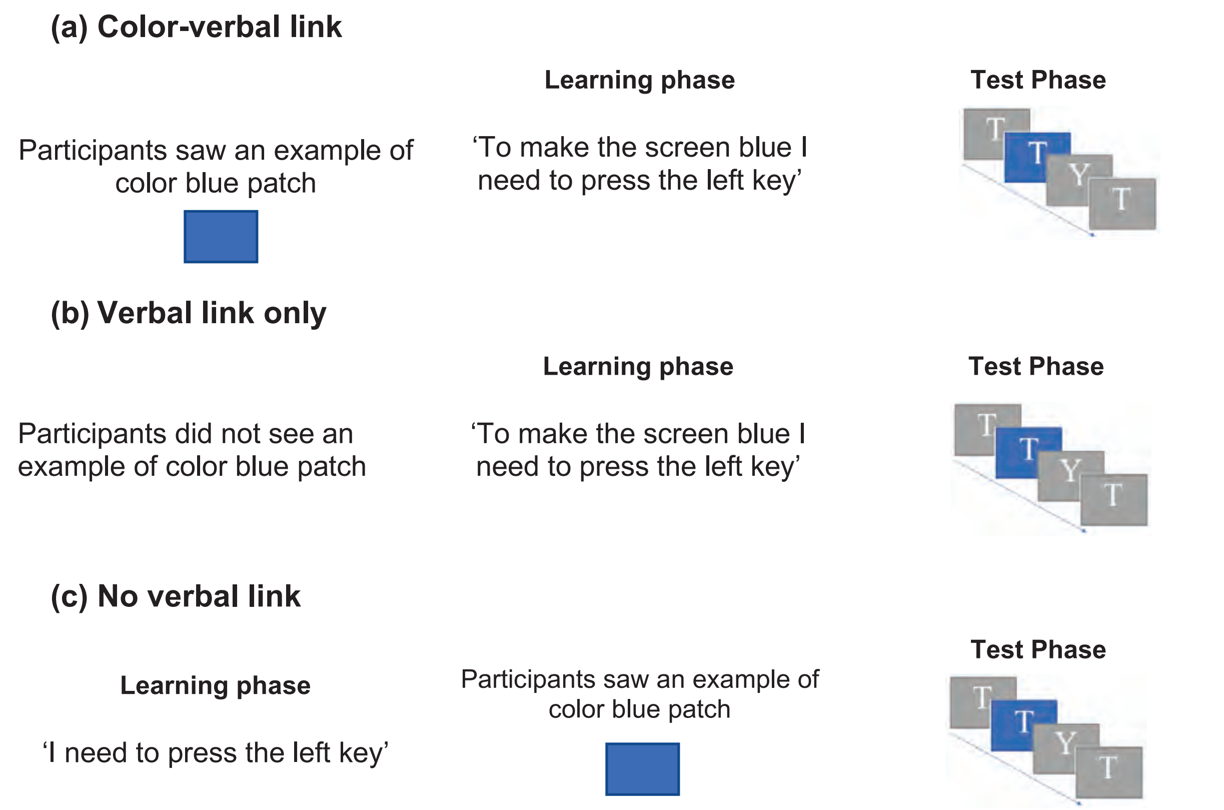
Most of our daily actions are aimed at achieving desired outcomes. Action-effect learning is based on the theoretical concept that actions are associated with their perceivable consequences through bidirectional associations (Elsner & Hommel, 2001). Past research has mostly investigated how these bidirectional associations are formed through active behavior using the two-stage model with learning and test phases (Elsner & Hommel, 2001; Waszak et al., 2012). In three studies, using the same two-stage model, we investigated the idea that verbally-induced action-effect associations can also result in action-effect associations even though participants never performed the behavior actively. In the first two online studies (*N* = 43, *N* = 400), participants memorized a specific action-effect instruction in the learning phase and completed a speeded categorization task in the test phase (Figure 1a). We assessed the consequences of the instructions by presenting the instructed effect as an irrelevant stimulus (blue background) in the reaction-time vowel-consonant classification task and compared response errors and response times for instruction-compatible and instruction-incompatible responses. Overall, in terms of response errors, we found evidence that verbal action-effect instructions led to associations between an action and perception (effect) that were automatically activated upon encountering the verbally-presented effect. Studies 1 and 2 are published in *Psychological Research* (Damanskyy et al., 2022).

In Study 3 (*N* = 700), we replicated the findings of Studies 1 and 2 and added two conditions to extend our theoretical understanding of verbal action-effect instructions and address potential methodological limitations of Studies 1 and 2. First, in Studies 1 and 2, participants saw a visual example of a critical stimulus specified in the action-effect sentence. Therefore, in Condition 2 (Figure 1b), we evaluated whether this visual presentation is a necessary precondition to observe a compatibility effect of verbally induced action-effect associations by eliminating the visual cue from the learning phase. In Condition 3 (Figure 1c), we evaluated whether making the critical stimulus and response salient could cause a compatibility effect. Therefore, the critical sentence included only information about the response (press left/right), and the critical stimulus was presented as a visual example *after* the critical sentence. Our results showed that the verbal action-effect sentence's compatibility effect was observed in Conditions 1 and 2 but not in Condition 3. Overall, the present studies provide empirical evidence that action-effect associations can be formed through verbal instructions. Although the perception-action relation presented as action-effect instructions was never executed by the participants before, it still had unintentional consequences when the perception component (effect) was encountered in the instruction-irrelevant classification task.

Keywords: action-effect, verbal instructions, action control, associative learning, language

Figure 1

A visual illustration of the experimental procedure of all three studies.



Note. Study 1 and 2 included only color verbal link condition (a). Study 3 included all conditions. In condition *a* and *c* participants were presented with an example of critical stimulus (color blue patch) that was specified in action-effect instructions in the learning phase. In condition *c* the example of irrelevant stimulus (color blue patch) was presented after the learning phase, and the verbal sentence in the learning phase did not include any specification of stimulus. The test phase included a vowel-consonant categorization task, in which participants had to categorize a presented stimulus as a vowel or consonant. In critical trials (i.e., trials with blue background) a target stimulus was accompanied by an irrelevant stimulus that served as the response effect in condition *a* and *b.* In control trials the irrelevant stimulus was absent (i.e., trials with gray background).

References

Damanskyy, Y., Martiny-Huenger, T., & Parks-Stamm, E. J. (2022). Unintentional response priming from verbal action–effect instructions. *Psychological Research.* https://doi.org/10.1007/s00426-022-01664-0

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 229-240. https://doi.org/10.1037/0096-1523.27.1.229

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: Neurophysiological basis and functional consequences. *Neurosci Biobehav Rev*, *36*(2), 943-959. https://doi.org/10.1016/j.neubiorev.2011.11.004

Incremental Negation Processing With Positive Questions under Discussion

Oksana Tsaregorodtseva, Elena Albu & Barbara Kaup (University of Tübingen)

Introduction. Negative sentences are generally harder to process compared to affirmative sentences (for an overview, see Kaup & Dudschig, 2020). However, the processing difficulty can be modulated by context. When used without any contextual support, negative sentences (*John hasn't washed the car*) are processed in a two-step fashion (Kaup et al., 2006): comprehenders first represent the non-factual object state (clean car) and only subsequently the factual object state (dirty car). One possible explanation is that, in the absence of any contextual information, negative sentences tend to answer a positive question under discussion (QUD) (e.g., Has John washed the car?). In contrast, negation is processed incrementally when used in a context in which it addresses a negative QUD (Tian et al., 2010; Wang et al., 2021), expressed either by means of cleft sentences (*It was John who didn't cook the spaghetti*) or wh- questions (*Which fruit isn't peeled?*).

Research question and predictions. In two behavioral experiments, we investigated whether negative sentences can be processed incrementally when addressing a positive polar QUD. To that end, in a probe recognition task (Fig. 1), affirmative and negative sentences (Table 1) were used in the absence (Exp 1) of in a supportive context (Exp 2) generated by discourse markers (*As expected, John has/hasn't washed the car*). These markers render the contextual expectation salient (It was expected that John would/ wouldn't) wash the car), restricting the set of possible QUDs to a positive one (Has John washed the car?). For affirmative sentences, we expect faster response times (RTs) for the factual compared to the non-factual object states in both experiments. In contrast, for negative sentences we expect different patterns of responses. If negation is processed in a two-step way, we expect a crossover interaction between the factors *Polarity (aff/neg)* and *Depicted object state (factual/non-factual)*, with faster RTs for the non-factual (clean car) than for the factual object state (dirty car). If negation is processed incrementally, we expect a main effect of *Depicted object state*, with faster RTs for the factual (dirty car) compared to the non-factual object state (clean car).

Results. RTs on correct picture-present trials were analyzed by means of a linear mixed effects model (Table 2). As predicted, participants responded faster to the factual object states in the affirmative condition in both experiments, while different patterns of responses emerged in the negative condition. *Experiment 1* (N = 104; 20 men; M_{age} = 37.26; *SD* = 12.39) revealed a crossover interaction between the factors *Polarity* and *Depicted object state* ($\chi 2(1) = 9.50$, $p = .002$, $\beta = -21.02$, $t = -3.08$), a main effect of Polarity ($\chi 2(1) = 10.36$, $p = .001$, $\beta = -21.93$, $t = -3.22$) but no main effect of Depicted object state ($\chi 2(1) = 2.47$, $p = .116$, $\beta = -10.75$, $t = -1.58$). In line with the two-step procedure, these findings suggest that participants responded faster to non-factual (clean car) than to the factual object states (Fig. 2a). By contrast, *Experiment 2* (N = 88; 27 men; M_{age} = 39.82; *SD* = 13.95) showed the reversed pattern with two main effects of Depicted object state ($\chi 2(1) = 23.77$, $p < .001$, $\beta = -41.78$, $t = -4.90$) and Polarity ($\chi 2(1) = 16.74$, $p < .001$, $\beta = -35.07$, $t = -4.11$). There was a significant ordinal interaction this time ($\chi 2(1) = 11.03$, $p = .001$, $\beta = 28.39$, $t = -3.32$), which suggests that participants responded faster to the factual (dirty car) compared to the non-factual object state in the negative condition (Fig. 2b). To receive more information about the pattern of responses in the negative conditions, a post hoc test was performed. This showed an interaction between the *Depicted object state* and *Context* (no context Exp. 1/context Exp. 2) ($\chi 2(1) = 4.51$, $p = .034$, $\beta = 16.84$, $t = 2.12$), which replicates previous findings according to which the processing of negative sentences is strongly modulated by context (Fig. 2c).

Discussion. All in all, the present paper corroborates previous results which indicate that context strongly influences the processing of negative sentences. Furthermore, it provides new evidence showing that, in a supportive context, negation can be incrementally processed when it addresses a positive polar QUD.

Keywords: negation processing, grounded cognition, pragmatics, language comprehension

Table 1. Example items for the probe recognition task in Experiments 1 & 2.

Polarity	Depicted object state	Experiments 1 & 2 Sentence	Display
Affirmative	factual	(As expected) John has washed the car.	
	non-factual	(As expected) John has washed the car.	
Negative	factual	(As expected) John hasn't washed the car.	
	non-factual	(As expected) John hasn't washed the car.	

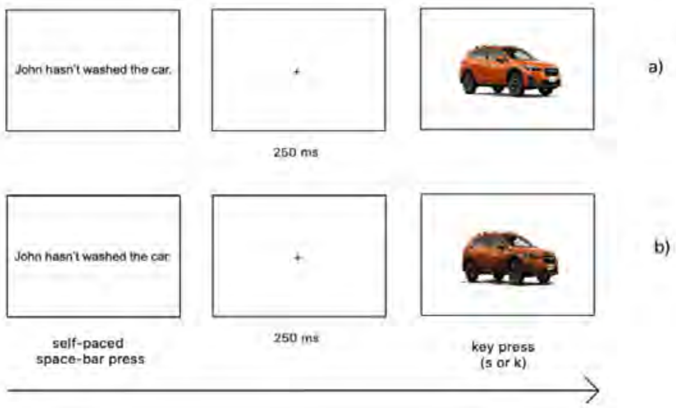


Figure 1. The time course of a typical negative trial; a) non-factual object state b) factual object state.

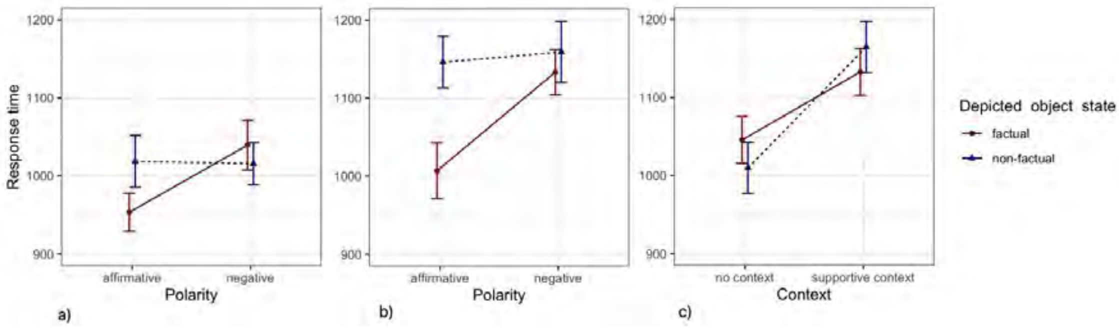


Figure 2. Response times Experiments 1 & 2, and Post hoc; error bars denote 95% confidence intervals; a) Experiment 1 b) Experiment 2 c) Post hoc.

Table 2. Linear mixed effects models in Experiments 1 & 2 and Post hoc.

Exp. 1 & 2	The base model: rt ~ Polarity + Depicted object state + (1| item + 1|participant)
	The best model: rt ~ Polarity* Depicted object state + (1| item + 1|participant)
Post hoc	The base model: rt ~ Context + Depicted object state + (1| item + 1|participant)
	The best model: rt ~ Context * Depicted object state + (1| item + 1|participant)

References

Kaup, B., & Dudschig, C. (2020). Understanding negation. In *The Oxford handbook of negation.*

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed?. *Journal of Pragmatics, 38*(7), 1033-1050.

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account.

Wang, S., Sun, C., Tian, Y., & Breheny, R. (2021). Verifying negative sentences. *Journal of Psycholinguistic Research, 50*(6), 1511-1534

Session 3

Gateway Identity and Spatial Remapping in a Combined Grid and Place Cell Attractor.

Tristan Baumann[1] and Hanspeter A. Mallot[1]

[1]Cognitive Neuroscience, Department of Biology, University of Tübingen
Auf der Morgenstelle 28, 72076 Tübingen, Germany

The representation of space in the rodent brain is generally attributed to place cells located in the hippocampus (O'Keefe & Dostrovsky, 1971; Moser et al., 2017). These cells cover the environment with location-specific firing fields ("place fields") and the population code uniquely describes each position in space. Place fields and the combined population code are subject to change if the rat enters a new compartment in the experimental maze. This effect, known as remapping (Muller & Kubie, 1987; Leutgeb et al., 2005), cannot be explained from path integration (grid cell activity) and local sensory cues alone but requires additional knowledge about the context. This is exemplified in two situations: Normally, when the animal returns to a known compartment, the place cells remap to the original pattern associated with that compartment. However, when the environment includes multiple visually identical but otherwise separate compartments, place cells also remap to the same pattern in each room (Grieves et al., 2016), and the animals confuse the rooms. The process specifically happens at the gateways or transitions, but not within a compartment, suggesting the reactivation of a stored pattern based on cues at the entrance.

We present a model for the hippocampal-entorhinal interplay in which the activity of place and grid cells follows a joint attractor dynamic. Place cells depend on the current grid cell activity but can also reset and change the grid cell activity in the remapping process. Remapping is triggered by the recognition of a gateway. When this happens, the previously stored pattern of place cell activity for this gateway is reactivated from a "gateway database". The joint attractor will then reinstate the grid cell pattern that was active when the gateway had first been learned and path integration can proceed from there. In this sense, the pattern of active cells provides both information about the current context and the position of the animal within that context. The model is tested with various mazes used in the experimental literature and reproduces the published results. We also make testable predictions for remapping in a new maze type. Based on the model, we hypothesize the existence of "gate cells" that drive the place cells and with them the joint hippocampal-entorhinal loop into the corresponding attractor whenever a gate is detected.

Keywords: attractors, remapping, place cells, grid cells, path integration

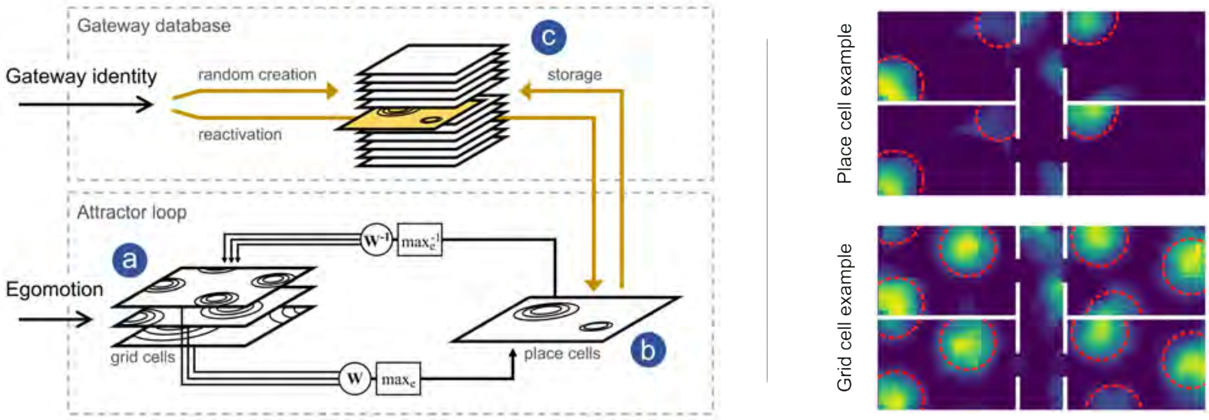


Figure 1: **Left:** Model overview. The grid cell module *(a)* consists of three separate attractor subnetworks that simulate grid cells at different scale and orientation. The combined activity is summed and filtered to form the place cell module *(b)*. The place cell module projects back to the grid cell module with the inverse weights to form a combined attractor. At gateways between compartments, the place cell pattern is stored and reactivated *(c)* to enable the model to return to a previous place cell pattern. **Right:** Examples of place and grid cell firing fields (red circles) in a multicompartment environment. The firing fields repeat in the identical parallel rooms (either left or right), but opposed rooms on either side of the central hallway are distinguished.

References

Grieves, R. M., Jenkins, B. W., Harland, B. C., Wood, E. R., & Dudchenko, P. A. (2016). Place field repetition and spatial learning in a multicompartment environment. *Hippocampus*, *26*(1), 118–134. doi: 10.1002/hipo.22496

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., & Moser, M.-B. (2005, July). Independent Codes for Spatial and Episodic Memory in Hippocampal Neuronal Ensembles. *Science*, *309*(5734), 619–623. doi: 10.1126/science.1114037

Moser, E. I., Moser, M.-B., & McNaughton, B. L. (2017, November). Spatial representation in the hippocampal formation: A history. *Nature Neuroscience*, *20*(11), 1448–1464. doi: 10.1038/nn.4653

Muller, R. U., & Kubie, J. L. (1987, July). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *Journal of Neuroscience*, *7*(7), 1951–1968. doi: 10.1523/JNEUROSCI.07-07-01951.1987

O'Keefe, J., & Dostrovsky, J. (1971, November). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. doi: 10.1016/0006-8993(71)90358-1

Learning Latent Event Codes for Hierarchical Prediction and Generalization

Christian Gumbsch[1,2], Georg Martius[2,*], & Martin V. Butz[1,*]
[*] equal supervision, [1] Neuro-Cognitive Modeling Group, University of Tübingen
[2] Autonomous Learning Group, Max Planck Institute for Intelligent Systems

Learning Sparsely Changing Event Codes

Artificial system can reach super-human performance for tasks they are trained on but often fail to generalize to minor task variations. Presumably because they lack compositional sensorimotor-grounded abstractions (Eppe et al., 2022). Humans seem to encode perceptual inputs by means of somewhat discrete, hierarchically organized event codes (Zacks et al., 2007; Butz et al., 2021). In contrast to the *discrete* event structure, most artificial systems, like recurrent neural networks (RNNs), *continuously* update their internal, or latent, representations.

We hypothesize that fostering event-encoding latent states in RNNs could improve generalization. For that, we propose Gated L_0-Regularized Dynamics (GateL0RD). GateL0RD (Fig. 1a) is an RNN that uses an update gate $\mathbf{\Lambda}_t \in [0, 1]$ to determine whether to update its latent state ($\mathbf{\Lambda}_t > 0$) or not ($\mathbf{\Lambda}_t = 0$). GateL0RD is trained using a special loss function: $\mathcal{L} = \mathcal{L}_{\text{task}} + \lambda \|\mathbf{\Lambda}\|_0$. Besides the task-based loss $\mathcal{L}_{\text{task}}$, an L_0-norm, weighted by a hyperparameter λ, punishes update gate openings. Thus, GateL0RD learns to only sparsely update its latent codes.

Robust Generalization in Prediction and Control

We first compared GateL0RD to other RNNs as the memory module in a forward model (Fig. 1b) in the Fetch Pick&Place (FPP) simulation. In FPP a robot arm can grasp and transport an object. We trained the systems to predict sequences of observations $\boldsymbol{o}_t$ containing rapid grasps of objects. We tested it on similar sequences (test) and ones where a grasp takes longer (generalization). All RNNs achieved low test prediction errors (Fig. 1c), but only GateL0RD achieved also a low generalization error (Fig. 1d). GateL0RD seems to overfit less to spurious temporal correlations by only encoding the reach- and grasp-events within its latent states (Fig. 1e).

We also used GateL0RD as a memory module in a reinforcement learning (RL) agent (Fig. 2a), replacing an RNN (vanilla). We applied it to problems of the MiniGrid gridworld (Fig. 2b), where GateL0RD not only learned faster to solve the problems (Fig. 2c), but also exhibited improved zero-shot generalization when applied to other variants of the problems during testing (Fig. 2d).

Learning Hierarchical Event-Based Predictions

Lastly, we learned hierarchical predictions via GateL0RD. To do so, we embedded GateL0RD in a forward-inverse model (Fig. 3a) and trained a feed-forward *skip network* to predict the observations for which GateL0RD's latent states change, i.e. the next event boundary. We trained the system on FPP sequences where the object is grasped, the hand points to the goal, or the arm is stretched. During testing, we then queried the skip network for hierarchical predictions. We show that the system learned to predict the approximate end point of where the gripper will move to, such as the location of an object when executing a reach (Fig. 3b).

Conclusion

We introduced GateL0RD, an RNN that develops sparsely changing, event-encoding latent states, akin to hypothesized encodings in humans. GateL0RD is better at generalization and learns temporal abstract predictions, which could enable flexible hierarchical problem solving.

Keywords: event cognition, recurrent neural networks, sparsity, generalization, control

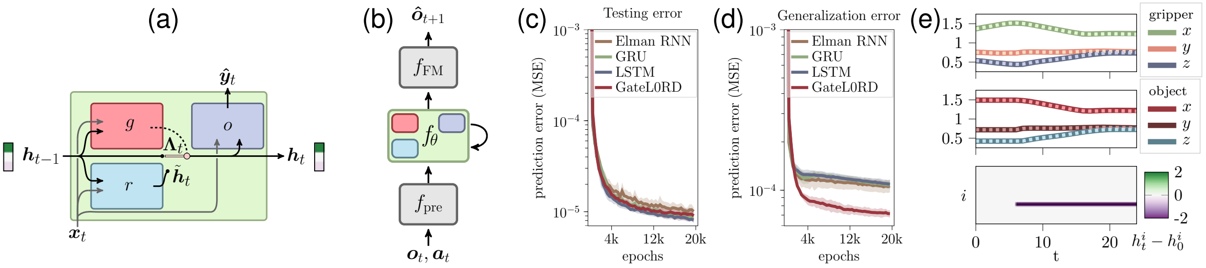


Figure 1: GateL0RD: (a) GateL0RD's structure, (b) forward model with feed-forward networks (FFN, grey) and an RNN (green). Fetch Pick&Place results for different RNNs: (c) test prediction error (grasp as in training), (d) generalization error (later grasp), and (d) exemplary grasping sequence with positions of gripper (top) and object (middle), and latent state $\boldsymbol{h}_t$ (bottom).

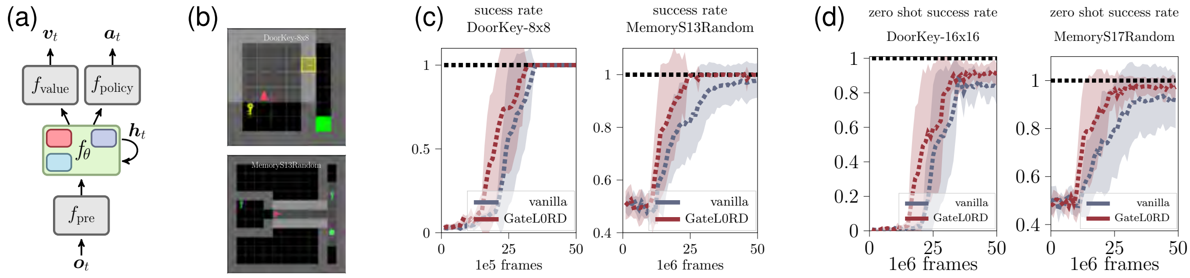


Figure 2: RL results: (a) RL architecture using an RNN (green; GateL0RD or Long Short-Term Memory (vanilla)), (b) MiniGrid (`github.com/maximecb/gym-minigrid`) tasks, success rate for (c) the trained task and (d) when tested on large versions of the tasks without retraining.

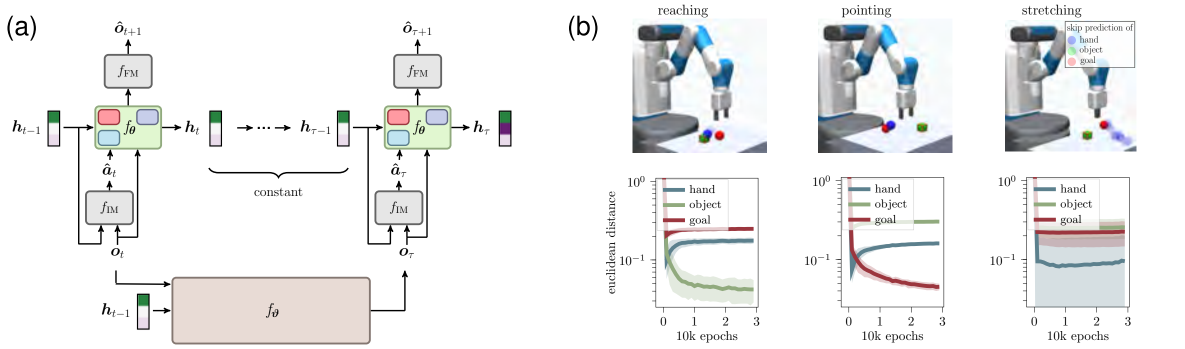


Figure 3: Hierarchical predictions: (a) the skip network $\boldsymbol{f}_{\vartheta}$ (bottom) is trained to predict the observations accompanying latent state changes of a forward-inverse model with GateL0RD (top, unrolled over time), (b) skip prediction at $t = 2$ for Fetch Pick&Place sequences showing exemplary predictions over 10 seeds (top), and the distances of skip-predicted position of the gripper to the current position of three entities (gripper, object, goal) (bottom).

References

Butz, M. V., Achimova, A., Bilkey, D., & Knott, A. (2021). Event-predictive cognition: A root for conceptual human thought. *Topics in Cognitive Science*, *13*(1), 10-24. doi: https://doi.org/10.1111/tops.12522

Eppe, M., Gumbsch, C., Kerzel, M., Nguyen, P. D., Butz, M. V., & Wermter, S. (2022). Intelligent problem-solving as integrated hierarchical reinforcement learning. *Nature Machine Intelligence*, 1–10.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological Bulletin*, *133*(2), 273-293. doi: 10.1037/0033-2909.133.2.273

Demands and potentials of different levels of neuro-cognitive models for human spatial cognition

Nicolas Kuske[1], Marco Ragni[2], Florian Röhrbein[1], Julien Vitay[1], and Fred Hamker[1]

[1]Department of Computer Science, Chemnitz University of Technology

[2]Department of Humanities and Social Sciences, Chemnitz University of Technology

The brain is composed of multiple overlapping networks competing and collaborating to implement various cognitive functions. Classically, models of human cognition have focused on planning processes which involve declarative knowledge [7]. Dual process theories purport two contrasting cognitive processes or systems [4]. In addition to declarative, planning (DP) processes, they consider procedural, habitual (PH) behavior. In the context of spatial cognition, navigating a well-known environment in which a path is blocked, would be a typical example of a DP process [8]. Encountering an intersection at which one regularly turns right on the shortest way to work, can trigger an automatic motor response even when the shortest way to the current goal location would require a left turn [11, 15]. Repeatedly rewarded actions lead to PH context-response associations.

We discuss and evaluate models of DP- and PH processes in spatial cognition on two different levels of organization. These are the biological implementation, and the symbolic algorithm level [6, 9, 12]. Arguably, the most prominent algorithmic model architecture is the Adaptive Control of Thought-Rational (ACT-R) [1, 13]. It is Turing complete and can emulate any cognitive feat which can be expressed through computational means. This expressive power, however, is also one of ACT-R's great drawbacks. While it does mimic the cognitive bottleneck appearing during serial planning processes, for example, ACT-R is generally underconstrained [14]. Seemingly, this algorithmic approach to understand cognition abstracted away too much biological detail. At the implementation level, the hippocampus is the brain region most prominently involved in DP processes [2, 8]. The basal ganglia make up an important hub shaping the PH system [11, 16]. We review recent neurocomputational models of spatial cognition. Moreover, we discuss the extent to which these biological models can provide natural boundary conditions for algorithmic models. As an exemplary use case, we compare the organization of cognitive learning processes at both model levels.

Reinforcement learning is a type of machine learning which closely resembles reward-related learning processes in animals and humans [10]. The prevalent theoretical perspective was that DP- and PH processes can be mapped onto two different classes of reinforcement learning algorithms [3]. Researchers distinguished between a model-based- and a model-free approach to reward-guided behavior. Recently, experimental and neurocomputational modeling work has shown that reinforcement learning using successor representations provides a neurobiologically more adequate description of reward-related learning processes in spatial cognition. [5, 17]. The cognitive model system ACT-R can be programmed to learn DP processes via a model-based reinforcement algorithm and PH processes through a model-free algorithm, independently [13]. The biological implementation model level underlying the successor representation algorithm, however, demands dependence of DP- and PH processes. We found a natural constraint for ACT-R's framework, exemplifying the potential of inter-model organization level comparisons.

Keywords: Dual Process Theory, Spatial Cognition, Cognitive Model Level, Reinforcement Learning

References

[1] Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

[2] Bellmund, J. L. S., Gärdenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. *Science*, *362*(6415), eaat6766. https://doi.org/10.1126/science.aat6766

[3] Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

[4] Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278.

[5] Geerts, J. P., Chersi, F., Stachenfeld, K. L., & Burgess, N. (2020). A general model of hippocampal and dorsal striatal learning and decision making. *Proceedings of the National Academy of Sciences*, *117*(49), 31427–31437. https://doi.org/10.1073/pnas.2007981117

[6] Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational Use of Cognitive Resources: Levels of Analysis Between the Computational and the Algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. https://doi.org/10.1111/tops.12142

[7] Kotseruba, I., & Tsotsos, J. K. (2020). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*, *53*(1), 17–94.

[8] Maguire, E. A., Burgess, N., Donnett, J. G., Frackowiak, R. S. J., Frith, C. D., & O'Keefe, J. (1998). Knowing Where and Getting There: A Human Navigation Network. *Science*, *280*(5365), 921–924.

[9] Marr, D., & Poggio, T. (1976). From understanding computation to understanding neural circuitry. *MIT Technical Report*. https://doi.org/1721.1/5782

[10] Neftci, E. O., & Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence*, *1*(3), 133–143. https://doi.org/10.1038/s42256-019-0025-4

[11] Packard, M. G., & McGaugh, J. L. (1996). Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. *Neurobiology of Learning and Memory*, *65*(1), 65–72. https://doi.org/10.1006/nlme.1996.0007

[12] Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science*. MIT Press.

[13] Ritter, F. E., Tehranchi, F., & Oury, J. D. (2019). ACT-R: A cognitive architecture for modeling cognition. *WIREs Cognitive Science*, *10*(3). https://doi.org/10.1002/wcs.1488

[14] Schultheis, H. (2009). Computational and explanatory power of cognitive architectures: The case of act-r. *Proceedings of the 9th international conference on cognitive modeling (iccm'09)*.

[15] Siegel, A. W., & White, S. H. (1975). The Development of Spatial Representations of Large-Scale Environments. *Advances in Child Development and Behavior* (pp. 9–55). Elsevier. https://doi.org/10.1016/S0065-2407(08)60007-5

[16] Smith, K. S., & Graybiel, A. M. (2013). A Dual Operator View of Habitual Behavior Reflecting Cortical and Striatal Dynamics. *Neuron*, *79*(2), 361–374. https://doi.org/10.1016/j.neuron.2013.05.038

[17] Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

A neural dynamic process model of scene representation, categorical visual search and scene grammar in natural scenes

Raul Grieben (raul.grieben@ini.rub.de)
Gregor Schöner (gregor.schoener@ini.rub.de)
Ruhr-Universität Bochum, Institut für Neuroinformatik
Universitätsstraße 150, 44801 Bochum, Germany

The world we live in is highly structured, and that structure induces expectations that strongly influence how we search and ultimately interact with objects in our environment. Yet the vast majority of studies in visual perception use artificial visual scenes and simplified stimuli. These greatly advanced our understanding of the basic principles underlying visual search, but provided limited understanding of how humans search for real-world objects in natural scenes. Wolfe (2021) proposed the concept of *scene guidance* and a special case of it has been formalized through the notion of *scene grammar* (Võ, 2021) based on experimental findings that *anchor objects* and their reproducible spatial relation to other objects enable humans to strongly reduce the area scanned in visual search. Alongside the increased interest in theories and models of *scene guidance* in the psychological domain, attention has also become an important topic in deep learning, although there is a considerable gap between the understanding of *attention* in these two fields.

We presented a neural process model for guided visual search and scene memory that did not only account for classical findings like feature vs. conjunction search, but also proposed answers to long-standing questions in the field of visual search: The influence of scene memory in the preview paradigm (Grieben et al., 2020) and the relationship between attention and feature binding (Grieben & Schöner, 2021) in the context of the unexpectedly efficient triple conjunction search. This model was limited to classical laboratory stimuli, however.

Here we present a neural process model that substantially extends Grieben et al. (2020) and Grieben & Schöner (2021) to natural scenes in a neurally plausible way. At the same time we incorporate a new neural process account for *scene grammar*. To our knowledge this is the first model to account for it. The neural process model is based on Dynamic Field Theory (DFT; Schöner et al., 2016) a mathematical framework that characterizes graded activation patterns of neural populations that evolve continuously in time. Enabling the model to interact with natural scenes required a major innovation, interfacing for the first time, a neural architecture based on DFT with a pre-trained headless deep convolutional neural network (CNN) that provides feature extraction. We use the biologically plausible BCM rule to map the distributed representation of the CNN feature maps to the localist representation of a 3D neural field defined over space and object identity. This enables the model to learn new object concepts from single exposures in a *supervised one-shot continual online learning* fashion. This interface makes it possible to combine the strength of the two frameworks. DFT delivers autonomous process organization, sequence generation, and working memory. CNNs are undoubtedly able to extract the complex features needed for object recognition. To guide visual search in natural scenes we also had to solve the problem of learning the association between categorical concepts and their corresponding preattentive shape guidance features. Our solution is to extract these guidance features from an intermediate layer of the CNN.

Keywords: neural dynamic process model; natural scenes; scene grammar; online learning; convolutional neural network.

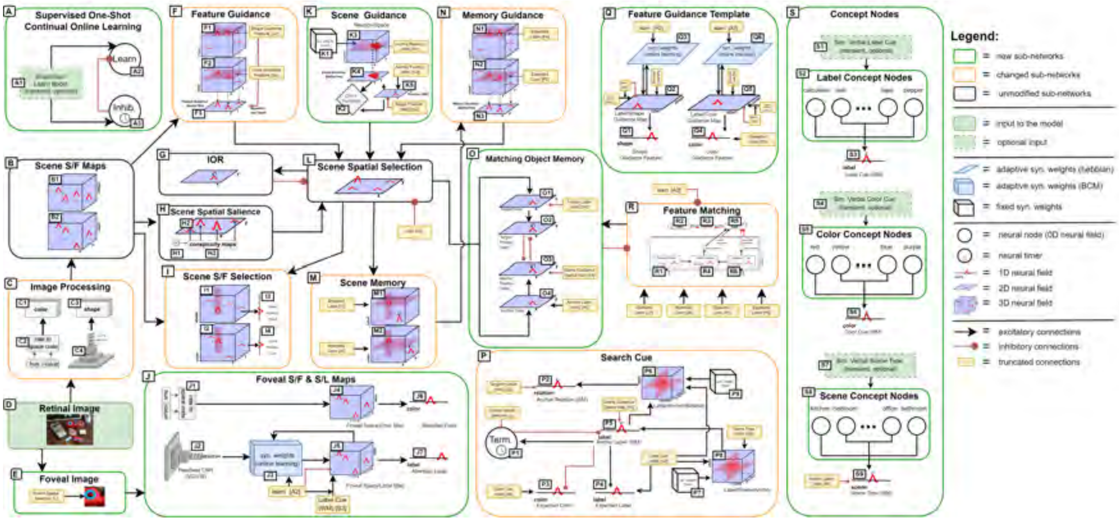


Figure 1: An overview of the neural dynamic process model.

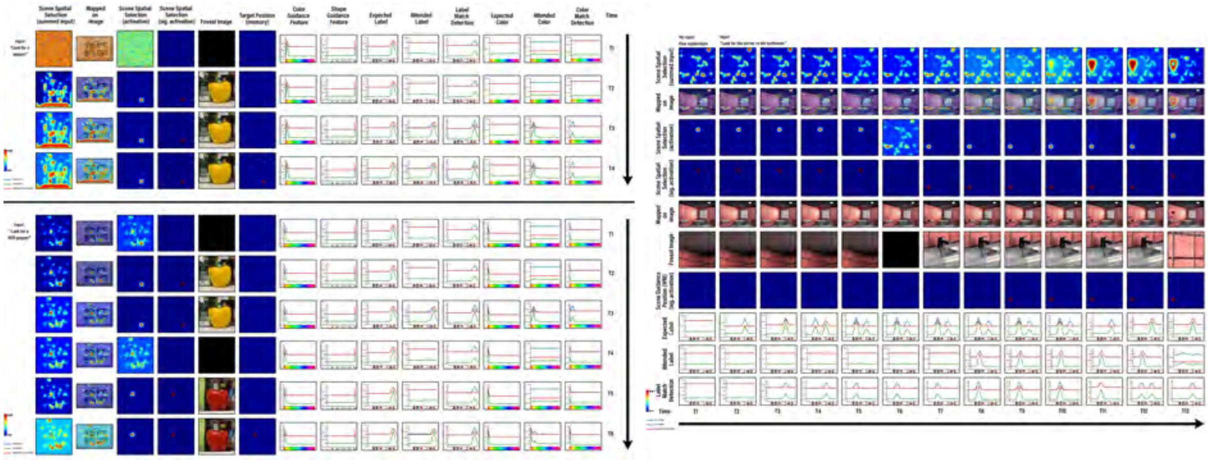


Figure 2: Left: Demonstration of the model performing two visual search tasks, with (bottom) or without (top) an extra color cue. Right: Demonstration of the model performing a scene guidance/grammar task (looking for the invisible mirror). Task image adapted from Võ (2021) (Fig. 5).

References

Grieben, R., & Schöner, G. (2021). A neural dynamic process model of combined bottom-up and top-down guidance in triple conjunction visual search. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd annual conference of the cognitive science society.* Cognitive Science Society.

Grieben, R., Tekülve, J., Zibner, S. K., Lins, J., Schneegans, S., & Schöner, G. (2020). Scene memory and spatial inhibition in visual search: A neural dynamic process model and new experimental evidence. *Attention, Perception, & Psychophysics*. doi: 10.3758/s13414-019-01898-y

Schöner, G., Spencer, J. P., & DFT Research Group, T. (2016). *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press.

Võ, M. L.-H. (2021). The meaning and structure of scenes. *Vision Research*, *181*, 10–20.

Wolfe, J. M. (2021). Guided search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, 1–33.

A neural dynamic model of action parsing

Minseok Kang[1,2,*], Jan Tekülve[2], and Gregor Schöner[2]

[1]Universität Osnabrück, Germany
[2]Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany
[*]Corresponding author: Minseok Kang, mikang@uos.de

When humans describe dynamic scenes unfolding in time, they do so in terms of discrete events such as "the boy kicked the ball to the goal" and "the girl caught the ball". Zacks et al. (2007) claim that such discrete events are not just reflections of our linguistic structures, but form fundamental psychological units that help us understand dynamic scenes, just as an object is a fundamental unit for understanding a visual scene. This study is concerned with how dynamic visual scenes are parsed and represented as sequences of observed actions, construed here a specific form of events.

The study presents a neural dynamic model of action parsing based on Dynamic Field Theory (DFT) (Schöner & Spencer, 2016). DFT is a mathematical theory that describes the temporal evolution of activation of neural populations in terms of dynamical systems. Neural fields are defined over feature dimensions that specific populations of neurons are tuned to. Localized patterns of supra-threshold activation are the units of representation in DFT. They arise and decay autonomously in response to changes in dynamic regimes that may be brought about by inputs. Cognitive neural architectures may be built by coupling multiple neural fields.

Figure 1 schematically illustrates the state of the model given an exemplary input provided as video clips of moving colored objects. Three aspects of action parsing are addressed by the model. First, the model generates descriptions of discrete actions involving objects based on object properties, relations between objects, and their change. Each description entails four components: an action verb, an agent object, a patient object, and a location object (Pastra & Aloimonos, 2012). A unique discrete node is activated for each component concept at the end of a demonstrated action. The set of activated concept nodes represent the parsed action so that different combinations of objects and action verbs is described by a finite set of nodes.

Second, relevant changes in the states of objects to be described should be selected while disregarding other aspects of the scene. The agent object is trivially identified by a unique color code. The location and patient objects are identified based on inputs from the finite set of relation detectors. Each relation detector is activated if the stipulated relationship between objects is fulfilled. Action verbs are differentiated based on activated relation detectors. The relations are encoded in patterns of heterogeneous synaptic connectivity (Lipinski et al., 2012).

Third, parsing dynamical scenes into discrete actions requires detecting their temporal boundaries within the video stream (Zacks et al., 2007). Each action verb has associated start, intermediate, and end conditions, which are represented by activating matching neural nodes. Whenever an end condition node is activated, the selected action verbs and objects activate a concept node that represents the parsed action. This entails extracting the feature value of each object (here, hue) and filling in the object's role.

By developing a neural dynamic model that parses sequences of actions from a visual stream, the study contributes to understanding how conceptual representations of actions may arise from the neural activation dynamics that are driven by sensory inputs.

Keywords: neural modelling, cognitive modelling, dynamical approach, event perception

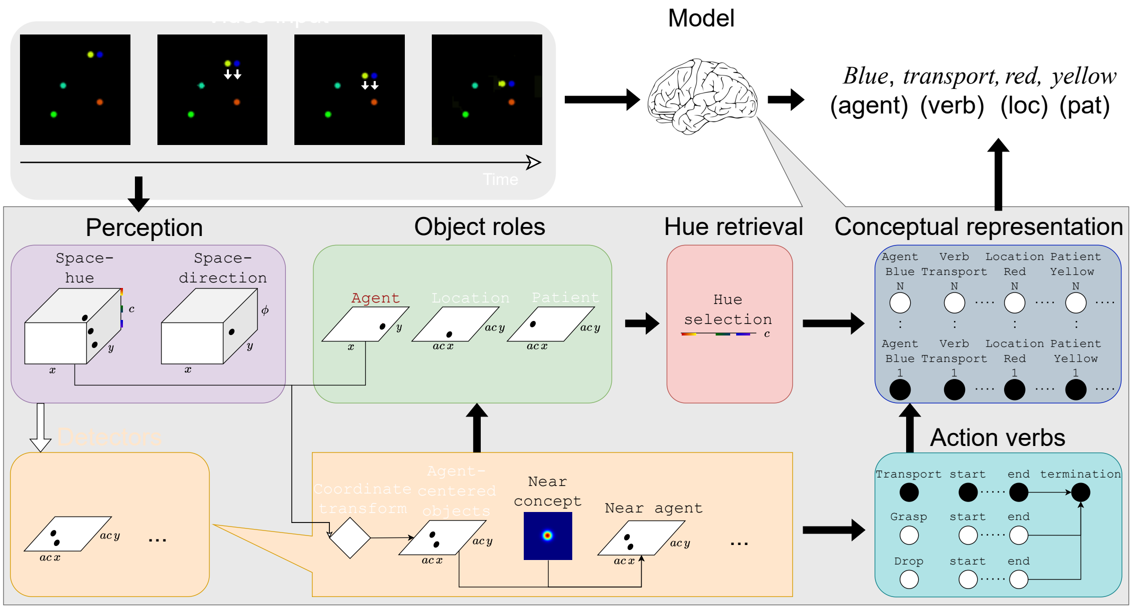


Figure 1: Overview of the neural field architecture. The input to the architecture is depicted on top left, and the interpretation of its conceptual representation is depicted on top right. Different sub-networks are highlighted according to their main functionality. Black dots within fields represent supra-threshold activations. Not all neural fields are shown and their connections are omitted for comprehensibility. As an example how the dynamics of connected neural fields give rise to cognitive functions, a detailed view of the near detector is depicted.

References

Lipinski, J., Schneegans, S., Sandamirskaya, Y., Spencer, J. P., & Schöner, G. (2012). A Neurobehavioral Model of Flexible Spatial Language Behaviors. *Journal of experimental psychology. Learning, memory, and cognition*, *38*(6), 1490–1511.

Pastra, K., & Aloimonos, Y. (2012). The minimalist grammar of action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585), 103–117.

Schöner, G., & Spencer, J. P. (2016). *Dynamic thinking: A primer on dynamic field theory*. Oxford University Press.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event Perception: A Mind/Brain Perspective. *Psychological bulletin*, *133*(2), 273–293.

An Embodied, Perceptually Grounded Neural Process Model Identifies Analogies In Accord With Structure Mapping Theory

Matthis Hesse, Daniel Sabinasz, Gregor Schöner
Institute for Neural Computation at Ruhr-University Bochum

Which neural processes underwrite the human competence to identify analogical structures? A common task that taps into that competence is to provide participants with a "base scene" and a "target scene" and ask them to find objects A and B in the base scene and objects C and D in the target scene such that A is to B as C is to D (see Figure 1).

According to Structure Mapping Theory (SMT; Gentner, 1983), a high-level theory of how humans identify an analogical mapping from base scene objects to target scene objects, only the relations between objects in a scene and relations between their features are relevant, and an optimal mapping is one which preserves as many relations as possible.

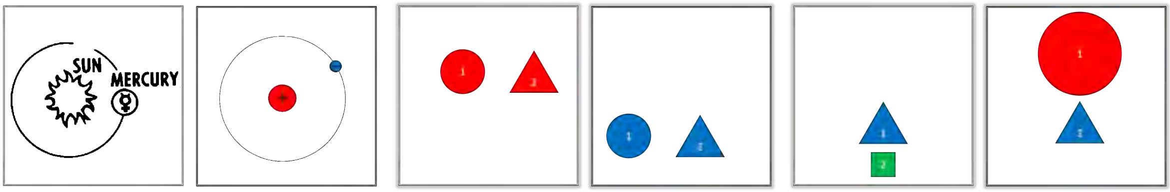
In line with the grounded cognition stance (e.g., Barsalou, 2008), we contend that analogy supervenes on the brain's modal systems for perception and action rather than on detached symbol manipulation. This is given support by the finding that people are more successful at identifying analogies if the analogous objects are superficially similar (Gentner & Maravilla, 2018), and by the observation that participants may be distracted into mapping objects based on superficial similarity (Figure 1c; Loewenstein & Gentner, 2005; Richland et al., 2006). These findings can be accounted for by assuming that perceptual representations of the to-be-mapped objects are taken into account.

Further, experimental findings demonstrate that people perform better at finding analogies when they know terms for the spatial relations and features (Richland et al., 2006; Christie & Gentner, 2014). Additionally, language cues may promote structure mapping over mapping based on superficial similarity (Loewenstein & Gentner, 2005). These considerations suggest that a conceptual description of the base scene underlies analogical mapping.

Driven by these considerations, we propose a dynamical systems model of the neural processes that can identify analogical mappings. It is presented with a base scene and extracts a conceptual description (e.g., "small red circle left of small red triangle; different shape, same color, same size"; Figure 1b). That conceptual description then guides a search for objects in a target scene which serve as candidates for analog objects, influenced by superficial similarity and fitting relational roles. After selecting a candidate mapping, it is evaluated according to the principles of SMT. Components of the conceptual description that do not match are removed, yielding a common description of the two scenes (e.g., "small circle left of small triangle; different shape, same color, same size"; Figure 1b). Our model is based on Dynamic Field Theory (Schöner et al., 2015) and uses established mechanisms for biased attentional selection, relation detection, hypothesis testing, and serial order processing.

We show the model's capabilities through simulations that demonstrate different qualitative behaviors depending on whether or not an analogy exists, and whether or not superficial similarity is helpful or distracting in the identification of an analogy. Figure 2 shows exemplary activation snapshots as the model finds the analogical mapping for the example in Figure 1b.

Keywords: analogy, dynamic field theory, neural process model, grounded cognition, embodied cognition



(a) *(b)* *(c)*

Figure 1: Three pairs of base and target scene. **(a)** Given the solar system as base scene and an atom as target scene, mercury is to the sun as the electron is to the nucleus. Importantly, the correct mapping cannot be found by comparing objects based on superficial similarity, but it has to be identified that the planets bear to the sun the same relationships (*orbiting*, *smaller than*) as the electrons to the nucleus. **(b)** The analogy consists in 1 bearing to 2 the same relationships (*left of*, *different shape*, *same color*, *same size*) in both scenes. **(c)** Mapping based on superficial similarity would incorrectly map base scene object 1 to target scene object 2, whereas structure mapping would map 1 to 1 and 2 to 2 because 1 bears to 2 the same relationships in both scenes (*above*, *different shape*, *different color*, *larger than*).

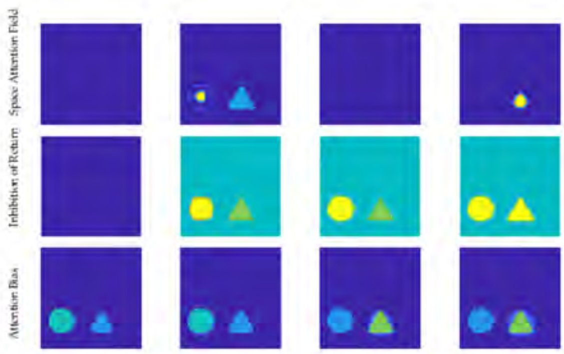


Figure 2: Visualization of our architecture as it identifies the analogical mapping in the target scene of Figure 1b. Snapshots show the activation of fields at one time. High activation is shown by a brighter color. The last row shows the sum of all attention bias contributions with a different color scale.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617-645.

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, *38*(2), 383-397.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy*. *Cognitive Science*, *7*(2), 155-170.

Gentner, D., & Maravilla, F. (2018). Analogical reasoning. In L. Ball & V. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 186–203). Routledege.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*(4), 315-353.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249-273.

Schöner, G., Spencer, J., & Research Group, D. F. T. (2015). *Dynamic thinking: A primer on dynamic field theory*. New York: Oxford University Press.

Session 4

The Dynamics of Body Ownership: Extending the Bayesian Causal Inference of Body Ownership Model Across Time

Moritz Schubert, Dominik Endres
Philipps University of Marburg

The rubber hand illusion (RHI) is one of the most common paradigms to study body ownership (Botvinick & Cohen, 1998). During an RHI, a rubber hand, lying in front of the participant, and their hidden hand are stroked in synchrony with a brush by the experimenter. Using this setup, a majority of participants experiences a illusionary ownership over the rubber hand.

According to Samad et al.'s (2015) model of the RHI (see fig. 1), participants do inference about the causal structure of the experiment based on the sensory input. If they conclude that the different sources of sensory input originate from the rubber hand as a common cause, they integrate said sensory input on the rubber hand and often feel like the rubber hand is their own. This illusion is not experienced if the sensory signals are too discrepant in time (or space). The model categorizes the vision of the rubber hand and proprioception of the real hand as *spatial* and the visual and tactile input resulting from the brush strokes as *temporal*. That said, there is no *actual* temporal component to the model: The priors are only updated once.

Schubert & Endres (2021) have pointed out that the width of the priors employed in Samad et al.'s (2015) model is unimaginably large (10^{35}) and hence psychologically implausible. They also demonstrated that when the prior widths are reduced to psychologically plausible values, the model does not yield ownership predictions in line with experimental data. Specifically, the model shows a high ownership posterior only for sensory discrepancies that are much smaller than those commonly used in RHI experiments.

We therefore hypothesize that a causal inference model might be able to reproduce experimental results if it is equipped with a component that accumulates sensory evidence across time. Presumably, after some time, enough evidence will have accumulated to overcome the discrepant spatial evidence indicating that the rubber hand and the real hand are located at different positions.

We have implemented a hidden Markov model (HMM) for sensory evidence accumulation across time. We calculated the posterior of the psychologically plausible *truncated model* reported by Schubert & Endres (2021) with the temporal component removed and used it as the prior for our HMM. In other words, the HMM replaced the temporal component of Samad et al.'s (2015) model. We ran the resulting model for three different experimental conditions: Synchronous input, asynchronous input and random input. In the first two conditions, the brush strokes have a constant length (1 second). In the synchronous condition, the strokes are exactly parallel, while in the asynchronous condition one of the signals is shifted. The random input serves as a baseline and is not typically used in RHI experiments. Our model was able to replicate the empirically observed effect that larger asynchrony reduces body ownership: For small discrepancies (see fig. 2, left) the asynchronous condition results in a high probability of a common cause (i.e. a body ownership illusion), while for large discrepancies (see fig. 2, right) no illusion occurs. Furthermore, we were able to simulate the occurrence of a body ownership illusion for psychologically plausible parameter choices. A key limitation of our model is that it predicts the occurrence of a body ownership illusion faster than is reported in the literature (see vertical line in fig. 2). Adjusting the model accordingly is the current focus of our research.

Keywords: body ownership, rubber hand illusion, Bayesian causal inference

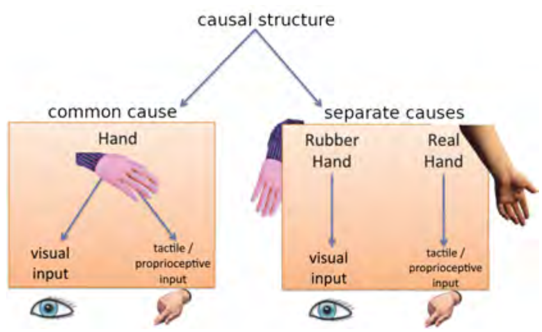


Figure 1: Graphical illustration of Samad et al.'s (2015) model. Modified version of Figure 1 in Samad et al. (2015), released under a Creative Commons Attributions License.

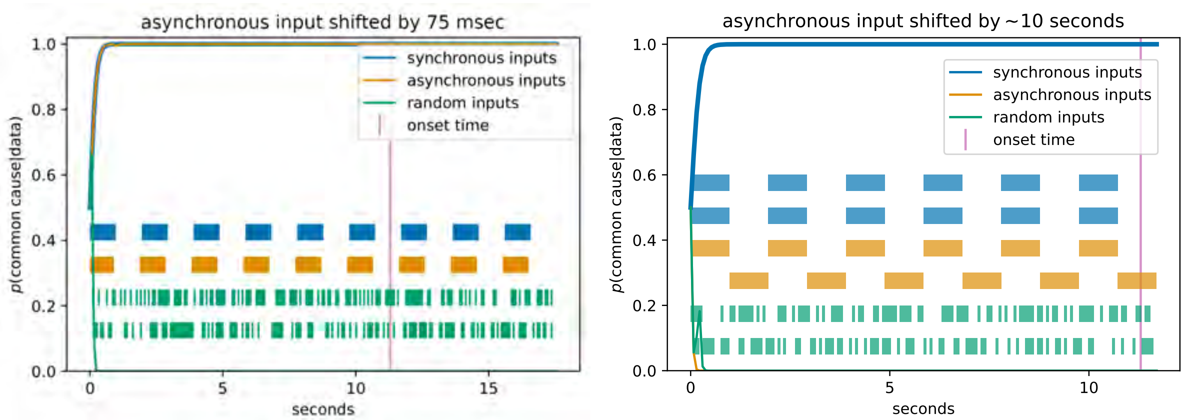


Figure 2: The lines indicate the posterior probability of a common cause across time. The vertical line indicates the expected onset time of the rubber hand illusion as reported by Ehrsson et al. (2004). The rectangles indicate the temporal sensory input in the different conditions with each rectangle representing one brush stroke. For each type of input, the upper row of rectangles indicates the visual input and the lower row the tactile input. **Left:** asynchronous input shifted by 75 msecs. **Right:** asynchronous input shifted by 975 msecs.

Data Availability Statement: The code used for generating the figures above can be found at `https://doi.org/10.5281/zenodo.6546881`

Acknowledgments: This work was supported by the IRTG 1901 "The Brain in Action", funded by the DFG, and "The Adaptive Mind", funded by the Excellence Program of the Hessian Ministry for Science and the Arts.

References

Botvinick, M., & Cohen, J. (1998, February). Rubber hands 'feel' touch that eyes see. *Nature*, *391*, 756. doi: 10.1038/35784

Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004, August). That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb. *Science*, *305*(5685), 875–877. doi: 10.1126/science.1097011

Samad, M., Chung, A. J., & Shams, L. (2015, February). Perception of Body Ownership Is Driven by Bayesian Sensory Inference. *PLOS ONE*, *10*(2), 1–23. doi: 10.1371/journal.pone.0117178

Schubert, M., & Endres, D. (2021, September). More Plausible Models of Body Ownership Could Benefit Virtual Reality Applications. *Computers*, *10*(9), 108. doi: 10.3390/computers10090108

Inference of Affordances and Active Motor Control in Simulated Agents

Fedor Scholz [1], Christian Gumbsch [1,2], Sebastian Otte [1] and Martin V. Butz [1]
[1]Neuro-Cognitive Modeling Group, Department of Computer Science and Department of Psychology, Eberhard Karls University of Tübingen, Tübingen, Germany
[2]Autonomous Learning Group, Max Planck Institute for Intelligent Systems, Tübingen, Germany

Flexible, goal-directed behavior is a fundamental aspect of human life. Based on the free energy minimization principle, the theory of active inference formalizes the generation of such behavior from a computational neuroscience perspective (Friston et al., 2015). In accordance with the theory, we introduce an output-probabilistic, temporally predictive, modular artificial neural network architecture (see Figure 1 (B)). Our system processes sensorimotor information, infers behavior-relevant aspects of its world, and invokes highly flexible, goal-directed behavior. We show that our architecture, which is trained end-to-end to minimize an approximation of free energy, develops latent states that can be interpreted as affordance codes (Gibson, 1986). That is, the emerging latent states signal which actions lead to which effects dependent on the local context. In combination with active inference and the cross-entropy method (Rubinstein, 1999), we show that flexible, goal-directed behavior can be invoked, incorporating the emerging affordance maps. We evaluate our architecture in an environment where a rocket-like agent can fly around by adjusting its throttles (see Figure 1 (A)). Our simulated agent flexibly steers through continuous spaces, avoids collisions with obstacles, and prefers pathways that lead to the goal with high certainty. An evaluation of the learned affordance codes (see Figure 1 (C)) indicates that the different sides of the obstacles are encoded similar to the corresponding sides of the environment's boundary, confirming behavior relevant encodings of the visually perceived environment. Additionally, we show that the learned agent is highly suitable for zero-shot generalization across environments: After training the agent in a handful of fixed environments with obstacles and other terrains affecting its behavior, it performs similarly well in procedurally generated environments containing different amounts of obstacles and terrains of various sizes at different locations (see Figure 1 (D)). In contrast, classical reinforcement learning agents often struggle with offline learning and generalization to similar environments. Furthermore, while certainly possible, it is not straight-forward how the look-up mechanism that is necessary for the emergence of affordance maps could be implement there. To improve and focus model learning further, we plan to invoke active inference-based, information-gain-oriented behavior also while learning the temporally predictive model itself in the near future. We expect that the addition of such self-motivated, curiosity-driven exploration of the environment will lead to even more stable, conceptual affordance-like encodings. By introducing architectural biases derived from event-segmentation theory, we intend to foster the development of both deeper event-predictive abstractions and compact, habitual behavioral primitives to enable deeper planning and reasoning. In sum, our system, as a cognitive architecture, offers a computational model of how affordance-like encodings may develop in our minds and how these encodings facilitate flexible, goal-directed behavioral adaptations under related but novel environmental circumstances.

Keywords: affordances, active inference, goal-directed control, simulation, free energy principle

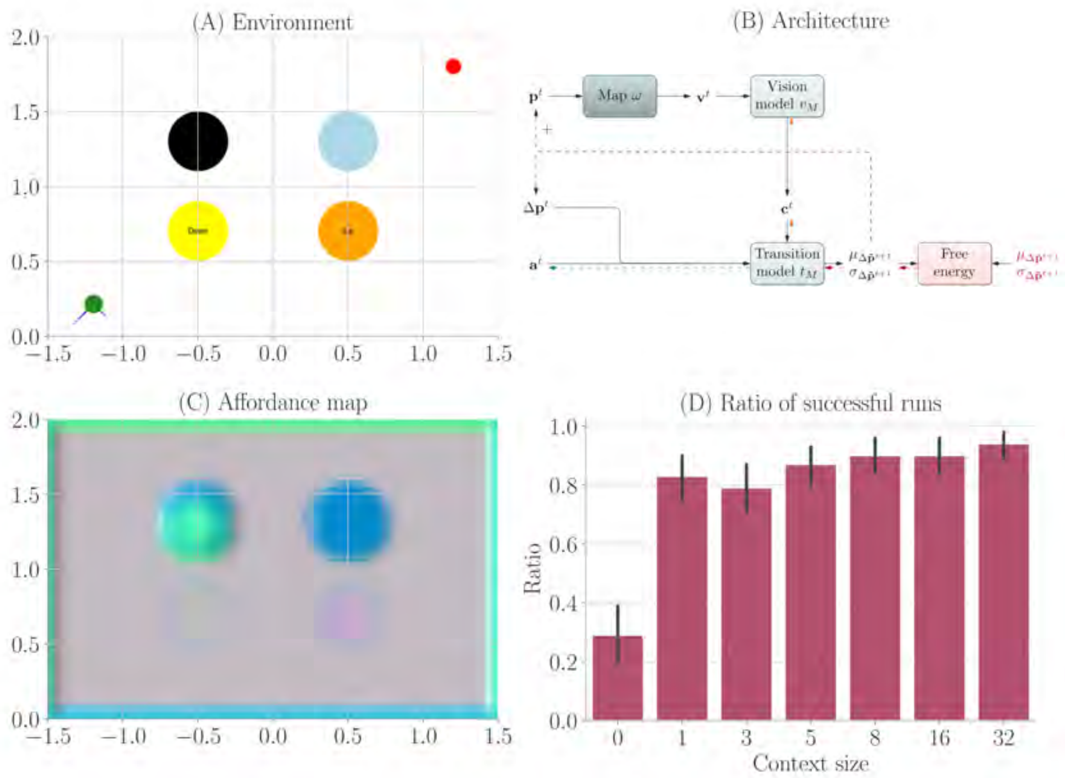


Figure 1: (A) Exemplary environment. The agent's (green) goal is to fly to the target (red) by adjusting its throttles (blue). Within the environment, obstacles (black), terrains of uncertainty (light blue), and force fields (yellow and orange) can have an effect on the agent's behaviour. (B) Architecture proposed in our paper. Based on the current position $\mathbf{p}^t$, the architecture performs a look-up in an environmental map ω. The vision model v_M receives the resulting visual information $\mathbf{v}^t$ and produces a contextual code $\mathbf{c}^t$. The transition model t_M utilizes this context $\mathbf{c}^t$, the last change in the position $\Delta\mathbf{p}^t$, and an action $\mathbf{a}^t$ to predict a probability distribution over the next change in position. During training, the loss between predicted and actual change in position is backpropagated onto t_M (red arrows) and further onto v_M (orange arrows) to train both models end-to-end. During planning, the map look-up is performed using position predictions. (C) Affordance map for exemplary environment with context size 16. The corresponding environment has an obstacle in the top left, a terrain of uncertainty in the top right, a force field pointing downwards in the bottom left, and a force field pointing upwards in the bottom right. To generate this map, we probed the environmental map at every sensible location, applied the vision model to each output, performed PCA to reduce the dimensionality to 3, and interpreted the results as RGB values. (D) Ratio of successful runs in procedurally generated environments with black lines representing standard deviations. A run was successful if the agent was closer to the target than 0.1 units in any time step and did not touch a terrain of uncertainty.

References

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*, 187–214.

Gibson, J. J. (1986). *The ecological approach to visual perception* (Vol. 1) [book]. Psychology Press New York.

Rubinstein, R. (1999). The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability*, *1*(2), 127–190.

The many faces of uncertainty in social interaction: a hierarchical model of metacognitive regulation in belief coordination

Sebastian Kahl & Stefan Kopp
Social Cognitive Systems Group, CITEC, Bielefeld University, Bielefeld

Social interaction is prone to misunderstandings, yet, its reciprocal processes enable robust recovery and allow for successful belief coordination, i.e., making sure you and your interaction partner know what you are talking about.

We present a hierarchy of nested generative sensorimotor perception and action processes based on the free-energy principle and active inference (Rao & Ballard, 1999; Clark, 2013; Friston et al., 2010; Kilner et al., 2007), which sees action as a form of inference over possible ways to make the environment meet the agent's predictions. The environment is affected through action to reduce uncertainty about predictions that stem from beliefs about the world (Adams et al., 2012). This has been discussed as a form of so-called *affordance competition*, a possible mechanism for action selection (Cisek, 2007; Pezzulo & Cisek, 2016) based on possible goals achievable through action. These principles are used to simulate the interplay of core processes within the so-called social brain (Van Overwalle, 2009), i.e., identified sensorimotor processing and mentalizing areas in the human brain, associated with activity during social interaction. Building on recent work on computational models of agency, sensorimotor- and mentalizing processes (Kahl et al., 2022; Kahl & Kopp, 2018; Pöppel et al., 2021), we propose a computational hierarchy that contributes a *bootstrapped* account of mentalizing. This enables a process for belief coordination similar to the *we-mode* proposed by Frith (2012): An implicit form of mentalizing that bootstraps the attribution of mental states that seems automatic, where contextual information and prior information can influence behavioral understanding top-down.

We argue that nested high-level representations that enable belief coordination do not only act to minimize uncertainty during perception and action, but they also assume a form of metacognitive regulatory control based on the uncertainty detected in the social interaction at large. That means, they can orchestrate the perception and production of social acts and are vital in resolving uncertainty in belief coordination (Fernandez-Duque et al., 2000). In other words, prediction errors during inference of social information may lead to repairs and other reciprocal communication behavior, which may be integral to communication in general (Healey et al., 2018). Technically, the presented hierarchy performs a general belief-update at each level, based on a linear-dynamic as well as an empirical Bayesian process that involves a dynamic information gain (or precision weighting) based in the system's uncertainty. This precision weighting can be biased by more strategically placed metacognitive regulation mechanisms to allow for weighted information gain to affect the agent's perception on a range from a focused attention to detail to a form of confirmation bias which ignores detailed prediction errors.

First, we introduce the foundational modeling approach and then show simulation results between three agents in a leader-follower non-verbal communication game. The agents perform reciprocal belief coordination that involves regulation of the whole nested hierarchy, its sensorimotor and mentalizing parts. The results stress the importance of balancing prior beliefs with new information by means of a precision-weighting bias that regulates not only perception and action, but also more generally, the gain on social information.

Keywords: social interaction, belief coordination, active inference, uncertainty, metacognition

References

Adams, R. A., Shipp, S., & Friston, K. J. (2012, November). Predictions not commands: Active inference in the motor system. *Brain Struct. Funct.*, *218*(3), 611–643. doi: 10/f4wkqx

Cisek, P. (2007, September). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philos. Trans. R. Soc. B Biol. Sci.*, *362*(1485), 1585–1599. doi: 10/ftxsfc

Clark, A. (2013, June). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.*, *36*(3), 181–204. doi: 10/f4xkv5

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000, June). Executive Attention and Metacognitive Regulation. *Consciousness and Cognition*, *9*(2), 288–307. doi: 10.1006/ccog.2000.0447

Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010, February). Action and behavior: A free-energy formulation. *Biol. Cybern.*, *102*(3), 227–260. doi: 10/c8zhrp

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, *367*(1599), 2213-2223. doi: 10/gddc2b

Healey, P. G. T., Mills, G. J., Eshghi, A., & Howes, C. (2018). Running repairs: Coordinating meaning in dialogue. *Top. Cogn. Sci.*, *10*(2), 367-388. doi: 10/gd8tpw

Kahl, S., & Kopp, S. (2018, December). A predictive processing model of perception and action for self-other distinction. *Front. Psychol.*, *9*, 47–14. doi: 10/gfr26d

Kahl, S., Wiese, S., Russwinkel, N., & Kopp, S. (2022, March). Towards autonomous artificial agents with an active self: Modeling sense of control in situated action. *Cognitive Systems Research*, *72*, 50–62. doi: 10/gpdzst

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007, April). Predictive coding: An account of the mirror neuron system. *Cogn. Process.*, *8*(3), 159–166. doi: 10/d66rpk

Pezzulo, G., & Cisek, P. (2016, June). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Sciences*, *20*(6), 414–424. doi: 10/f8pb23

Pöppel, J., Kahl, S., & Kopp, S. (2021, November). Resonating Minds—Emergent Collaboration Through Hierarchical Active Inference. *Cognitive Computation*. doi: 10/gpdzs7

Rao, R. P., & Ballard, D. H. (1999, January). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.*, *2*(1), 79–87. doi: 10/drddxm

Van Overwalle, F. (2009, March). Social cognition and the brain: A meta-analysis. *Hum. Brain Mapp.*, *30*(3), 829–858. doi: 10/bqd7c7

Session 5

Asymmetric Cross-Recurrence Plots for correlating time series of different length

Sebastian Wallot[1], Flavia Felletti[1], & Henning Drews[2]

[1]Institute for Sustainability Education and Psychology, Leuphana University, Germany

[2]Section of Epidemiology, University of Copenhagen, Denmark

One challenge in quantifying similarities of - and differences between - two time series or sequences is that time-bound processes do not have the same duration. This is particularly prominent for naturally occurring phenomena, such as cultural practices or physiological processes – e.g., the comparison of scan paths during reading (von der Malsburg, 2015), of sleep stages profiles (Drews et al., 2020), or of texts of different kinds (Grabowski, 2008).

In order to correlate such data series, an equal amount of time-matched data points is needed for each of the two data series to be compared, and the standard approach to make data series conform to this requirement includes usually either trimming of the longer time series or procedures for stretching/compressing of the data series at hand. Stretching or compressing can be archived by different operations, such as re-sampling or interpolation. However, this presupposes that the implicit or explicit modeling assumptions involved in these procedures adequately reflect the dynamics of the data at hand. In other words, they presuppose a relatively good a priori understanding of the dynamics of data, which is often not available when such data are being investigated. In so far as the assumptions of these procedures do not meet the actual temporal structure of the data at hand, trimming, re-sampling, and interpolation procedures can lead to the introduction of spurious correlations or the removal of existing correlations. This problem is even more prominent for the analysis of nominal sequences, as re-sampling or interpolation procedures are not as easily applicable for this kind of data compared to continuously sampled data.

Here, we present Asymmetric Cross-Recurrence Quantification Analysis (AsCRQA) as method that offers one solution to this problem. AsCRQA is an extension of Cross-Recurrence Quantification Analysis (CRQA; Shockley et al., 2002), a nonlinear bivariate correlation technique to quantify correlation and coupling between two time series. Quantification of correlation in CRQA is done extracting features from a cross-recurrence plot, a square-matrix that charts repetitions between two time series (Figure 1c). AsCRQA extends CRQA by allowing to compute recurrence-based similarity measures from a rectangular matrix that allows for two time series to differ in length (Figure 1d). Moreover, AsCRQA is readily applicable to nominal sequences of different length. We will describe the method, apply it to nominal data series, and compare those results to results obtained by using linear correlations together with trimming and stretching/compression techniques, based on simulation data.

Figure 1a shows a classical, symmetric cross-recurrence plot of two time series with the same underlying deterministic dynamics that differ only in their implied noise component. Figure 1b shows an asymmetric cross-recurrence plot for two time series with the same deterministic dynamics, but one of them had been re-sampled in exponentially increasing intervals. This means, that the second time series is shorter, and not merely a linearly down sampled version of the longer one. Still, both cross-recurrence plots look qualitatively similar. Moreover, when comparing cross-recurrence measures to nominal sequence correlation of the two time series of different length when applying cutting or linear resampling approaches, it is apparent that asymmetric cross-recurrence plots provide a more consistent measure of similarity (see Table 1). That is, AsCRQA provides very similar correlation parameters for time series with similar dynamics, no matter whether the match or differ in length, while Cramer's V does not.

*Keywords: correlation analysis, cross-*recurrence quantification analysis, time series of different length.

References

Drews, H. J., Wallot, S., Brysch, P., Berger-Johannsen, H., Weinhold, S. L., Mitkidis, P., ... & Göder, R. (2020). Bed-sharing in couples is associated with increased and stabilized REM sleep and sleep-stage synchronization. *Frontiers in Psychiatry*, *11*, 583.

Grabowski, J. (2008). The internal structure of university students' keyboard skills. *Journal of Writing Research*, *1*, 27–52.

Shockley, K., Butwill, M., Zbilut, J. P., & Webber Jr, C. L. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters A*, *305*, 59-69.

von der Malsburg, T., Kliegl, R., & Vasishth, S. (2015). Determinants of scanpath regularity in reading. *Cognitive science*, *39*, 1675-1703.

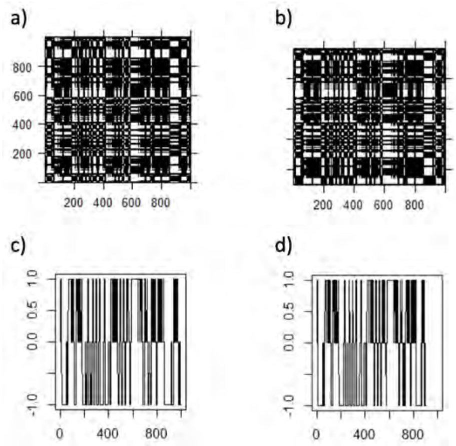


Figure 1. a) Classical (symmetric) cross-recurrence plot for time series of same length. b) Asymmetric cross-recurrence plot for time series of different length (c) and (d). c) Categorical time series of 1000 data points. d) Categorical time series as in (c), but down samples to only 900 data points by removing every 10th data point from (c).

Table 1: Comparison of Cross-Recurrence Measure LAM and Cramer's V.

Correlation measure	Timeseries of same length	Time series of different length
CRQA – LAM	0.222 [0.193, 0.253]	0.201 [0.174, 0.229]
Cramer's V	0.728 [0.704, 0.751]	N/A
Cramer's V (longer series trimmed to 900 data points at beginning)	N/A	0.046 [0.011, 0.091]
Cramer's V (longer series trimmed to 900 data points at end)	N/A	0.047 [0.011, 0.094]
Cramer's V (longer series linearly down sampled to 900 data points)	N/A	0.036 [0.020, 0.055]

Note. LAM = Laminarity Rate. A cross-recurrence measure of similarity of two time series; Cramer's V = correlation coefficient for two nominal series. Values of LAM and V in the column time series of same length provide something like a ground truth (i.e., the real association strength between the time series). The numbers in brackets provide the 95%-confidence intervals based on 1000 simulations. Values of LAM and V in the column time series of different length provide a measure of how the different approaches relate to the ground truth.

Deep neural networks as mechanistic explanations of object recognition – in search of the explanans

Bojana Grujičić
Max Planck School of Cognition
Humboldt-Universität zu Berlin, Berlin School of Mind and Brain
University College London, Science and Technology Studies Department

An array of recent results suggests that DNNs trained for an object recognition task are currently the best neuroscientific models for predicting the hierarchy of neural responses along the ventral stream (Lindsay, 2021; Kriegeskorte, 2015). This novel field aims to offer a new methodology for neuroscience in contrast to the traditional one (Nastase et al., 2020), having hopes of fulfilling not just its predictive but its explanatory goals as well (Lindsay, 2021; Cichy & Kaiser, 2019; Kietzmann et al., 2019; Kriegeskorte, 2015). In this paper, I inquire whether DNNs can be explanatory of a cognitive explanandum – that of the object recognition capacity in humans.

Beside this surge of interest in DNNs in neuroscience, philosophical interest in neural networks has been lately rising, with several arguments offered suggesting how neural network models could in principle explain mechanistically (Stinson, 2018) or that they are already mechanistic explanations of the object recognition capacity (Cao & Yamins, 2021; Buckner, 2018).

I analyse the claim of DNNs being mechanistic explanatory models of the object recognition capacity (Cao & Yamins, 2021; Buckner, 2018) and I inquire what exactly the DNN-based mechanistic explanans would be. Looking at the current research practice at the intersection of deep learning and neuroscience, I outline three different options. 1) Individual nodes in a DNN, their point-to-point connections and their organisation are the relevant entities, activities and organisational properties that map adequately onto the brain (Kaplan & Craver, 2011). 2) Neural manifolds in high-dimensional state spaces are basic entities and transformations over them are basic operations. 3) The variable of architecture of a DNN, which along with the objective function and the learning rule constrains the emergence of processes responsible for object recognition.

I will not try to arbitrate between these mechanistic explanantia of the object recognition capacity. Rather, I focus on the findings of Storrs et al. (2021) and Mehrer et al. (2020) in order to paint a picture of the variety of DNNs used in the field. Storrs et al. (2021) show that nine architecturally different convolutional DNNs all exhibit the hierarchical correspondence with the processing along the ventral stream (Storrs et al., 2021). Mehrer et al. (2020) analyse properties of the instances of architecturally identical DNNs trained from different random seeds and conclude that they differ. The recognition of this variety of DNNs prompts the question of the exact DNN-based mechanistic explanans of object recognition, as these networks seem to differ on all three above-mentioned accounts – on the level of individual nodes, neural manifolds and architectures. What is then a shared mechanism between these instances of DNNs that should map onto the brain's mechanism for object recognition? Are these convolutional DNNs nevertheless relevantly similar, are they relevantly dissimilar, or relevantly dissimilar but only exhibiting individual differences?

One way of uncovering shared features of DNNs relies on using similarity measures. I analyse different similarity measures in play, show that they are based on different assumptions about what are the relevant properties of neural patterns and show that some of them deliver opposing verdicts regarding the similarity of DNNs. Which similarity measure is the adequate one? I claim that this has to be arbitrated relative to the explanandum capacity. I concur with some recent calls to characterise the explanandum capacity of object recognition more elaborately (Lonnqvist et al., 2021; Ma & Peters, 2020), and I offer a methodological proposal to help arbitrate between similarity measures which relies on more extensively probing DNNs behaviourally and seeing in which cases judgements of similarity of processing across networks as indexed by a particular similarity measure covary with judgements of their similarity on the level of task performance. Because similarity measures themselves operate on different levels of analysis of a system, arbitrating between similarity measures will streamline the search for the explanans. Before we progress on this ground, I will claim, it is underdetermined which one (or more) of the three accounts of the DNN-based mechanistic explanans of the object recognition capacity is the adequate one.

Keywords: deep neural networks, object recognition, mechanisms, explanation, similarity measures

References

Buckner, C. (2018). Empiricism without magic: Transformational abstraction in deep convolutional neural networks. *Synthese*, *195*(12), 5339–5372.

Cao, R., & Yamins, D. (2021). Explanatory models in neuroscience: Part 1 - taking mechanistic abstraction seriously. *ArXiv:2104.01490*

Cichy, R. M., & Kaiser, D. (2019). Deep Neural Networks as Scientific Models. *Trends in Cognitive Sciences*, *23*(4), 305–317.

Kaplan, D. M., & Craver, C. F. (2011). The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective*. *Philosophy of Science*, *78*(4), 601–627

Kietzmann, T., McClure, P., & Kriegeskorte, N. Deep Neural Networks in Computational Neuroscience. *Oxford Research Encyclopedia of Neuroscience*

Kriegeskorte, N. (2015). Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science*, *1*(1), 417–446

Lindsay, G. W. (2021). Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future. *Journal of Cognitive Neuroscience*, *33*(10), 2017–2031

Lonnqvist, B., Bornet, A., Doerig, A., & Herzog, M. H. (2021). A comparative biology approach to DNN modeling of vision: A focus on differences, not similarities. *Journal of Vision*, *21*(10), 17.

Ma, W. J., & Peters, B. (2020). A neural network walks into a lab: Towards using deep nets as models for human behavior. *ArXiv:2005.02181*

Mehrer, J., Spoerer, C. J., Kriegeskorte, N., & Kietzmann, T. C. (2020). Individual differences among deep neural network models. *Nature Communications*, *11*(1), 5725

Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: Rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, *222*, 117254

Stinson, C. (2018). Explanation and connectionist models. In M. Sprevak & M. Colombo (Eds.), *The Routledge Handbook of the Computational Mind* (1st ed., pp. 120–133). Routledge.

Storrs, K. R., Kietzmann, T. C., Walther, A., Mehrer, J., & Kriegeskorte, N. (2021). Diverse Deep Neural Networks All Predict Human Inferior Temporal Cortex Well, After Training and Fitting. *Journal of Cognitive Neuroscience*, 1–21

Making Sense of the Natural Environment

Christoph von der Malsburg (FIAS, Frankfurt and INI, ETH Zürich), Benjamin Grewe (INI, ETH Zürich), Thilo Stadelmann (CAI, ZHAW Winterthur)

The neural basis of cognition is unclear to this day. We here present a conceptual framework resolving the conflict Fodor & Pylyshyn (1988); Dever (2006) between symbolic and neural approaches. In our scheme, the cortical carriers of meaning are not individual neurons but sets of neurons supporting each other by mutual excitation. These sets and their supporting connectivity are called 'net fragments' or simply 'fragments.' Also fragments activate only as part of larger nets composed of overlapping fragments. Fragments play the role of composite symbols. As each neuron can be part of several fragments, and each fragment can overlap with several alternative other fragments, fragments can be likened to jigsaw puzzle pieces that fit together in innumerable different arrangements. Any such arrangement must, however, conform to a highly non-trivial consistency condition.

Net fragments and the composite nets they form are supported by specific patterns of synaptic connections. These are formed in development and learning by network self-organization, a process studied experimentally Goodhill (2007) and theoretically Willshaw & von der Malsburg (1979); Häussler & von der Malsburg (1983) on the example of the ontogenetic establishment of retinotopic fiber projections. This process selects net structures that are sparse (limited fan-in and fan-out of connections at each neuron) and are self-consistent such that a sufficient number of fibers converge on any one neuron from within the net. The composition rule for fragments to co-activate in a net is that together they form a net that is self-consistent (and would be stable under the process of network self-organization). Any particular large net (that is, set of active neurons) is unlikely to occur more than once in a life-time, so that only relatively small fragments have a chance to be active again and again to thus reach stability under network self-organization. But as these fragments overlap in multiple ways, cortex develops into an overlay of net fragments that supports an infinitude of consistent large-scale nets.

Among all possible thus-defined net structures a particular role is played by those that realize schema application. Each schema is an abstract structural description under which large numbers of instances can be united Bartlett (1932); Minsky (1974); Schank & Abelson (1977). Invariant object recognition has been modeled as schema application Arathorn (2002); Olshausen et al. (1995); Hinton (1981); Kree & Zippelius (1988); von der Malsburg (1988) realizable as a net that is representing schema, instance and the structure-preserving mapping between them Wolfrum et al. (2008). Natural intelligence may be defined as the ability of pursuing vital goals and intentions in varying contexts. Behavioral control has been classically described as schema application Shettleworth (2010). We propose the composition of nets out of fragments as basis for this process von der Malsburg et al. (2022).

In distinction to present-day artificial neural networks the human brain can learn and generalize from very few examples. It is a well-established insight Geman et al. (1992); Wolpert (1996) that such efficiency must be based on a deep structural relationship between learning system and domain. Inherent in our neural representation framework is therefore the claim that also the environment can be seen as a composite of a finite set of structural fragment types.

Keywords: neural representation, network self-organization, compositionality, net fragments, behavioral schema, intentions.

References

Arathorn, D. (2002). *Map-seeking circuits in visual cognition – a computational mechanism for biological and machine vision*. Stanford, California: Standford Univ. Press.

Bartlett, F. (1932). *Remembering, a study in experimental and social psychology*. Cambridge: Cambridge University Press.

Dever, J. (2006). Compositionality. In E. Lepore & B. Smith (Eds.), *The oxford handbook of philosophy of language* (pp. 633–666). Oxford University Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1), 3-71. doi: 10.1016/0010-0277(88)90031-5

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, *4*, 1-58.

Goodhill, G. J. (2007). Contributions of theoetical modeling to the understanding of neural map development. *Neuron*, *56*, 301-311.

Häussler, A. F., & von der Malsburg, C. (1983). Development of retinotopic projections: An analytical treatment. *J. Theoretical Neurobiology*, *2*, 47–73. Retrieved from `https://vfs.fias.science/d/3cfce0fe5a/files/?p=/Retina.pdf`

Hinton, G. E. (1981). A Parallel Computation that Assigns Canonical Object-Based Frames of Reference. In *International joint conference on artificial intelligence* (pp. 683–685).

Kree, R., & Zippelius, A. (1988). Recognition of topological features of graphs and images in neural networks. *J. Phys. A*, *21*, 813-818.

Minsky, M. (1974, June). *A framework for representing knowledge* (Tech. Rep. No. 306). MIT AI Laboratory.

Olshausen, B., Anderson, C., & Van Essen, D. (1995). A multiscale dynamic routing circuit for forming size- and position-invariant object representations. *Journal of Computational Neuroscience*, *2*, 45-62.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. New Jersey: Erlbaum.

Shettleworth, S. (2010). *Cognition, evolution, and behavior (2nd ed.)*. Oxford: Oxford University Press.

von der Malsburg, C. (1988). Pattern recognition by labeled graph matching. *Neural Networks*, *1*, 141–148.

von der Malsburg, C., Stadelmann, T., & Grewe, B. (2022). *A theory of natural intelligence.* doi: 10.48550/ARXIV.2205.00002

Willshaw, D. J., & von der Malsburg, C. (1979). A marker induction mechanism for the establishment of ordered neural mappings; its application to the retinotectal problem. *Philosophical Transactions of the Royal Society of London, Series B*, *287*, 203–243.

Wolfrum, P., Wolff, C., Lücke, J., & von der Malsburg, C. (2008). A recurrent dynamic model for correspondence-based face recognition. *Journal of Vision*, *8*(7), 34. doi: 10.1167/8.7.34

Wolpert, D. (1996). The lack of a priori distinctions between learning algorithms. *Neural Computation*, *8*, 1341-1390.

Bayesian Cognitive Science Between Normativity and Descriptivity

Corina Strößner, Birkbeck, University of London
Ulrike Hahn, Birkbeck, University of London

Recent years have seen lively debate about different frameworks of cognition from both a descriptive (how do we think) and a normative (how should we think) basis. The relation between these spheres is complicated and the use of rational norms and optimality-based arguments, in particular within Bayesian cognitive science, has raised critical comments (Elqayam & Evans 2011, Bowers & Davis 2012, Jones & Love 2011). Hahn (2014) concludes that the debate requires a more fundamental approach. This talk offers such a fundamental analysis by employing concepts from philosophy of science, such as *stance*, *view*, *framework*, and *model*. The wider conceptual perspective on the debate will allow individual positions and the types of evidence that speak to them to be more clearly positioned. The talk will highlight two aspects that play a major role in the discussion: first the relation between normative and descriptive aspects within a framework, and second the general question of how to evaluate frameworks.

<u>*Frameworks and the interplay between normative and descriptive models*</u>:
A framework provides the means to formulate models. In case of the Bayesian framework, beliefs and information are translated to the probabilistic calculus. There are strong arguments in favour of the normative aptness of having beliefs that are probabilistically coherent. A Bayesian model (i.e., a particular probability distribution) is thus primarily a model of rational agents. The critical issue for Bayesian cognitive science is the relation between the model of the rational (or optimal) agent and a descriptive model. The most direct connection is to interpret a rational model as a candidate for a descriptive model and to test it (e.g., Hahn & Oaksford 2007). However, the normative model can also be used to explicate implicit assumptions that yield a rational explanation of data (e.g., the rarity assumption in Oaksford & Chater 1994). The legitimacy of the approach is based on the acceptance of the framework and its methodology (e.g., Bayesianism and rational analysis), which is by no means uncontroversial. This brings us to the second question: How can we evaluate frameworks?

<u>*Evaluating frameworks: stances and views*</u>
Since frameworks are not falsifiable, it is difficult to confirm or reject them (see Carnap 1950). In principle, there are two ways to approach frameworks, namely from the perspective of a stance and the perspective of a view. The concept of a stance (van Fraassen 2002) refers to the way to approach an issue and a style of reasoning (Rowbottom & Bueno 2011): The choice of a framework is "justified" by its fit to the epistemic attitude of a researcher and entails no commitments. Accordingly, the use of optimality-based arguments does not commit one to the thesis that human cognition is optimal. It rather means that the researcher favours a rational explanation in the style of Anderson (1991) over a mechanistic one.

Recently, Cuffaro & Hartmann (2021) contrasted a stance with a view. Like stances, views are connected to general attitudes towards the research subject and justify the choice of a framework. However, they are deeply connected to beliefs and are rationally justified. For instance, Bayesian cognitive scientists subscribe to the view that human cognition is concerned with uncertainty and that Bayesianism provides the appropriate solutions for dealing with uncertainty. In the talk, we will discuss how stance-like and view-like attitudes justify the choice of a framework for building a particular model that involves normative and descriptive aspects.

The talk will focus primarily on Bayesianism and its models, as these are the focus of current debates. For comparison, we will also look at other normative and descriptive frameworks and models from the field of reasoning, such as conceptual spaces (Osta-Vélez & Gärdenfors 2020), dynamic logic (van Benthem 2011), or mental model theory (Johnson-Laird 2010).

Keywords: Bayesianism, normative vs descriptive, rational analysis, philosophy of cognitive science

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological review*, 98(3), 471-85

van Benthem J. (2011) *Logical dynamics of information and interaction*. Cambridge University Press

Bowers, J. S. & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological bulletin*, 138(3), 389-414

Cuffaro, M. E., & Hartmann, S. (2021). The open systems view. arXiv preprint:2112.11095.

Carnap, R. (1950) Empiricism, Semantics, and Ontology, *Revue Internationale de Philosophie*, 4(11): 20–40.

Elqayam, S. & Evans, J. S. B. (2011). Subtracting "ought" from "is": Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, 34(5), 233-290.

van Fraassen, B. C. (2002). *The Empirical Stance.* Yale University Press.

Hahn, U. (2014). The Bayesian boom: good thing or bad? *Frontiers in psychology*, 5, 765.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*(3), 704–732

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243-18250.

Jones, M., & Love, B. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences, 34*(4), 169-188.

Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608-631.

Osta-Vélez, M. & Gärdenfors, P. (2020). Category-based induction in conceptual spaces. *Journal of Mathematical Psychology* 96, p. 102357.

Rowbottom, D. P., & Bueno, O. (2011). How to change it: modes of engagement, rationality and stance voluntarism. *Synthese*, 178(1), 7-17

Interfield Integration in Cognitive Science

Nina Poth
Department of Philosophy II
Ruhr-Universität Bochum

Doubts about the unity of cognitive science have recently been raised. While some argue that the field is disunified because it lacks true interdisciplinarity, integration, & intellectual coherence to a high degree (Núñez et al. 2019), respondents argue that its multidisciplinary character illustrates no threat to progress in the field (Gentner 2019; Bender 2019; French 2019). I focus on two fundamental questions to guide this debate: What would unity in cognitive science amount to in the first place & in what sense is it valuable to making progress in the field?

My claim is that the unity of cognitive science should be understood in terms of *interfield integrations*, which allow researchers from different fields to share solutions to a problem by transforming proper terms (Darden & Maull, 1977). For example, 'mutation' has been transformed from genetics, where it means 'heritable alteration in the genotype of an organism', to biochemistry, where it is understood as 'heritable alteration in base sequence'. The key advantage of unification in this sense is that it *expands scientific knowledge* by strengthening semantic connections across subfields: rather than being replaced, "claims about mutation from genetics [are] retained & biochemical claims added" (Maull, 1977, p. 152).[1] To avoid conflating similar but distinct problems across fields, Darden & Maull distinguish successful & unsuccessful cases. The former 'shift' a problem to a new field context, maintaining the knowledge associated with its original description. The latter merely 'import' the problem, ignoring how it arose in the initial field context. Importations likely produce misunderstanding & lack of differentiation, thereby diminishing the chances to expand knowledge across field boundaries. On this view, investigating the unificatory status of cognitive science amounts to searching for interfield theories.

I apply this view to investigate a possible integration of psychology, philosophy & AI in the study of concept learning. The theory of Bayesian concept learning (Ullman & Tenenbaum 2020) explicates the implicit or vague relationships among these fields by showing how the cognitive function of concepts as studied by developmental psychology & philosophy can be investigated in terms of the computational structure of hierarchical Bayes nets as studied in AI. While 'concept learning' in AI means 'correct classification of sparse data' (Lake et al. 2017), it refers to generalisation, bootstrapping & perceptual belief justification in psychology & philosophy. This integration is promising insofar as the problem of concept learning is shifted (not imported) from philosophy and psychology to AI. An open challenge for AI is to maintain an understanding of concept learning in terms of bootstrapping or model expansion (cf. Rutar et al. 2022).

Unity in cognitive science need not be associated with grand explanatory theories. Its value consists in expanding knowledge and establishing collaborations. Furthermore, the case of Bayesian concept learning illustrates that interfield integrations additionally allow researchers to share methodologies (e.g., Bayesian modelling), empirical evidence (e.g., on children's word

[1] Fields are defined as domains of facts that center on a focal problem. For instance, neuroscience centres on the problem of information-processing in the nervous system, psychology focuses on the problem of cognitive capacities & how they produce their behavioural effects, AI focuses on the problem of devising machines that learn & think. Which problem is focal to a field depends on researchers' explanatory goals & on the availability of specific tools & methods, which set expectations on how the problem should be solved.

learning abilities), and problem solutions. Indeed, nativist & rationalist ideologies in philosophy & psychology highlight the importance of innate constraints (Carey 2009; Margolis & Laurence 2011), suggesting to encode information about category variability in the priors of a generative model to solve the problem of few-shot learning (Lake et al. 2017; Smith et al. 2021). Finally, contrary to their classical opposition, this novel view on unification resonates with recent calls for pluralism in cognitive science (Potochnik & Oliveira 2020). For instance, AI reverse-engineering often focuses on accurate performance in comparison to humans, while experimental psychology typically focuses on infant development & armchair philosophy on conceptual foundations. They all study concept learning, but each focuses on different behavioural, statistical & conceptual patterns to do so. It is exactly because of this plurality of explanatory styles that a collaboration among these fields becomes relevant to a greater understanding of cognition.

Keywords: unification, cognitive science, interfield integration, concept learning

References

Bender, A. (2019). The value of diversity in cognitive science. *Topics in Cognitive Science*, *11*(4), 853–863.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Darden, L., & Maull, N. (1977). Interfield theories. *Philosophy of Science*, *44*(1), 43–64.

French, R. M. (2019). Missing the forest for the trees: Why cognitive science circa 2019 is alive and well. *Topics in Cognitive Science*, *11*(4), 880–883.

Gentner, D. (2019). Cognitive science is and should be pluralistic. *Topics in Cognitive Science*, *11*(4), 884–891.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and brain sciences*, *40*.

Margolis, E., & Laurence, S. (2011). Learning matters: The role of learning in concept acquisition. *Mind & Language*, *26*(5), 507–539.

Maull, N. L. (1977). Unifying science without reduction. *Studies in History and Philosophy of Science Part A*, *8*(2), 143–162.

Núñez, R., Allen, M., Gao, R., Miller Rigoli, C., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behaviour*, *3*(8), 782–791.

Potochnik, A., & Sanches de Oliveira, G. (2020). Patterns in cognitive phenomena and pluralism of explanatory styles. *Topics in Cognitive Science*, *12*(4), 1306–1320.

Rutar, D., de Wolff, E., van Rooij, I., & Kwisthout, J. (2022). Structure learning in predictive processing needs revision. *Computational Brain & Behavior*, *5*(2), 234–243.

Smith, R., Schwartenbeck, P., Parr, T., & Friston, K. J. (2020). An active inference approach to modeling structure learning: Concept learning as an example case. *Frontiers in Computational Neuroscience*, *14*, 41.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *The Annual Review of Developmental Psychology is*, *2*, 533–58.

Session 6

‘Active attention guidance’ as a basic scaffold for (social) cognition

Maja Griem, Ruhr-Universität Bochum

Abstract:

The environment shapes our behaviour. This idea has received a comeback during the past decades. However, the focus has mainly been on objects of the world, such as white canes, calculators or notebooks. Since our environment is made up of objects and other agents, I would like to shift the attention towards how other agents might shape our behaviour and facilitate our cognitive capacities.

Take pointing behaviour as an example: By pointing towards a third entity, I can intentionally direct the other’s attention towards a common target. This cannot only facilitate or enhance the other’s understanding of the content of my talk, but also help to shape the other’s attention in various ways. Furthermore, pointing seems to develop in parallel to the ability to establish joint attention in human infants (Cappuccio & Shepherd, 2013), thereby implying an important role of pointing within socio-cognitive development. Yet, pointing gestures are primarily understood as an extended finger used to point to a specific target (Krause, 2018; Heschl, 2018). This poses a problem if we include other cultures, that show a variety of different pointing signals, such as pointing with the whole hand, chin, nose, or other body parts (Wilkins, 2003), not to speak of other species that might even be physically unable to do so.

To account for these differences, we need to extend the notion of pointing by shifting the question from ‘how does the gesture look?’ to ‘in which way does it facilitate or enhance the other’s cognition?’ I propose a theory of minimal pointing that focuses on the underlying mechanisms, namely (1) attention gaining, and (2) directing the other’s attention towards a specific target. I call this *active attention guidance,* since it might not be as sophisticated as declarative forms of human pointing but requires more than passive gaze following: the subject looking at an object first has to actively catch and guide the other’s attention towards the target. This can be achieved by various means and cues (visual, auditory, etc.) that gain the other’s attention and direct it towards the specific target, requiring a sensitivity for the other’s attention without the need for highly sophisticated linguistic abilities. However, for a strategy to count as AAG, the effects must be actively pursued showing goal directed behaviour. Therefore, rigid patterns have to be ruled out by restricting AAG to behaviours allowing for at least minimal flexibility. Consequently, attention gaining and directing behaviours relying on a sensitivity for others’ attentional states and provide minimal flexibility and goal directedness qualify as AAG. I argue that AAG itself is the basis for joint attention as it includes the active guidance of the other’s attention, which is thought of as a major developmental step in human social cognition (Tomasello, 1999). Nonhuman animals can be included without the need to provide evidence for (explicit) general mental state ascription, as has classically been required for more sophisticated forms of joint attention. Instead, a sensitivity for others’ attentional states alone (i.e., without second or third order intentionality required) would be sufficient. Such a sensitivity has been shown in various species, such as apes (Dafreville et al., 2021), other primates (Hattori et al., 2010), cats (Ito et al., 2016), dogs (Miklósi et al., 2000; Call et al., 2003), wolves (Range & Virányi, 2011), horses (Trösch et al., 2019), and corvids (von Bayern et al., 2009). Figure 1 provides a simple version of an experimental setup to test for AAG in non-linguistic species by focusing on potential corrective behavior in an advanced gaze-following paradigm.

Further, I argue that *active attention guidance* can be seen as a basic scaffold for social cognition, underlying flexible coordination, cooperation, and other more sophisticated socio-cognitive abilities. Since our environment consists not only of objects, but also to a huge extent of other agents, we should put more emphasis on how others can shape our cognition and what we can do to help others. To achieve this, a broader comparative view including different species can be useful to examine more basic underlying mechanisms of highly developed cognitive abilities within an even more complex social environment.

Keywords: pointing, attention, scaffold, comparative cognition, social cognition

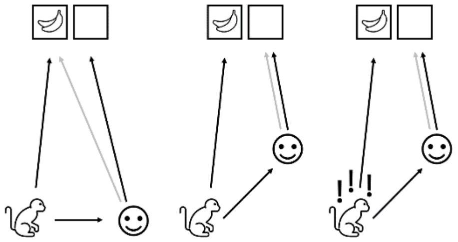


Figure 1 (from left to right): a) the gazer (monkey) tries to actively direct the attention of the follower to the box containing a banana. The long distance between subjects and objects as well as the small distance between the two boxes, causes an ambiguity in the gazes of gazer and follower. The black arrow indicates the follower's actual gaze direction, whereas the grey arrow is the gazer's potential misperception of the follower's gaze direction. b) The follower is now near enough by the objects for the gazer to clearly see that the follower is looking at the empty box. If the gazer really tries to actively direct the follower's attention, we expect the gazer to c) register the error of the follower and show a behavioral change, e.g., a higher frequency in gaze-alternation between the banana box and the follower or even a corrective behavior towards the follower, to make it attend the desired target.

References

Call, J., Bräuer, J., Kaminski, J., & Tomasello, M. (2003). Domestic dogs (Canis familiaris) are sensitive to the attentional state of humans. *Journal of Comparative Psychology*, 117(3), 257–263. https://doi.org/10.1037/0735-7036.117.3.257

Cappuccio, M. L., & Shepherd, S. V. (2013). Pointing Hand: Joint Attention and Embodied Symbols. In Z. Radman (Ed.), *The Hand, an Organ of the Mind: What the Manual Tells the Mental*. (pp. 303–326). MIT Press.

Dafreville, M., Hobaiter, C., Guidetti, M., Sillam-Dussès, D., & Bourjade, M. (2021). Sensitivity to the communicative partner's attentional state: A developmental study on mother–infant dyads in wild chimpanzees (Pan troglodytes schweinfurthii). *American Journal of Primatology*, 83(12). https://doi.org/10.1002/ajp.23339

Hattori, Y., Kuroshima, H., & Fujita, K. (2010). Tufted capuchin monkeys (Cebus apella) show understanding of human attentional states when requesting food held by a human. *Animal Cognition*, 13(1), 87–92. https://doi.org/10.1007/s10071-009-0248-6

Heschl, A. (2018). From grasping to pointing: The evolution of referentiality in man and animals. In E. M. Luef & M. M. Marin (Eds.), *The talking species: Perspectives on the evolutionary, neuronal and cultural foundations of language* (pp. 79–104). Uni-Press Graz Verlag GmbH.

Ito, Y., Watanabe, A., Takagi, S., Arahori, M., & Saito, A. (2016). Cats beg for food from the human who looks at and calls to them: Ability to understand humans' attentional states. *Psychologia*, 59(2–3), 112–120. https://doi.org/10.2117/psysoc.2016.112

Krause, M. A., Udell, M. A. R., Leavens, D. A., & Skopos, L. (2018). Animal pointing: Changing trends and findings from 30 years of research. *Journal of Comparative Psychology*, 132(3), 326–345. https://doi.org/10.1037/com0000125

Miklósi, A., Polgárdi, R., Topál, J., & Csányi, V. (2000). Intentional behaviour in dog-human communication: An experimental analysis of "showing" behaviour in the dog. *Animal Cognition*, 3(3), 159–166. https://doi.org/10.1007/s100710000072

Range, F., & Virányi, Z. (2011). Development of Gaze Following Abilities in Wolves (Canis Lupus). *PLoS ONE*, 6(2), e16888. https://doi.org/10.1371/journal.pone.0016888

Tomasello, M. (1999). *The cultural origins of human cognition.* Harvard University Press.

Trösch, M., Ringhofer, M., Yamamoto, S., Lemarchand, J., Parias, C., Lormant, F., & Lansade, L. (2019). Horses prefer to solicit a person who previously observed a food-hiding process to access this food: A possible indication of attentional state attribution. *Behavioural Processes*, 166, 103906. https://doi.org/10.1016/j.beproc.2019.103906

von Bayern, A. M. P., & Emery, N. J. (2009). Jackdaws Respond to Human Attentional States and Communicative Cues in Different Contexts. *Current Biology*, 19(7), 602–606. https://doi.org/10.1016/j.cub.2009.02.062

Wilkins, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 171–215). Lawrence Erlbaum Associates Publishers.

Trading off subjective values and social expectations for interpersonal trust

Gabriele Bellucci
Department of Computational Neuroscience
Max Planck Institute for biological Cybernetics

Trust is central to a large variety of social interactions. Different research fields have empirically and theoretically investigated trust, observing trusting behaviors in different situations and pinpointing its different components and constituents. However, a unifying, computational formalization of those diverse components and constituents of trust is still lacking. Previous work has mainly used computational models borrowed from other fields and developed for other purposes to explain trusting behaviors in empirical paradigms. In particular, most of those models propose that social behaviors are generated by subjective utility functions that trade off the expected benefits and costs of the prosocial act. However, these models overlook knowledge from social psychology suggesting that individuals weigh different social expectations (e.g., that others have of them and they themselves have of others) before engaging in a variety of social behaviors. Here, I propose that a computational model for social behaviors should combine current and prospective action values with social beliefs and expectancies about one's and a partner's behavior. Hence, the proposed model extends previous cost/utility functions by accounting for social expectations about a partner's likely, future behavior. I then showcase how this can be done by computationally formalizing a verbal model of trust, that is, the vulnerability model that defines trust as the willingness to accept vulnerability to others based on their trustworthiness. In the context of the classic investment game (IG)—an economic game thought to capture some important features of trusting behaviors in social interactions (Figure 1), I show how variations of a single parameter of the vulnerability model that governs the balance between a scalar subjective value and learnt social expectations, generates varying behaviors that can be interpreted as stemming from different "trust attitudes". I then show how these behavioral patterns change as a function of an individual's loss aversion and expectations of a partner's behavior, allowing for changes in trust contingent on the partner's trustworthiness like psychological models predict. I finally demonstrate how the vulnerability model can be easily extended in a novel IG paradigm to investigate different expectations about an interacting partner based on different types of information that contribute to expectations of the partner's trustworthiness. In particular, I will focus on benevolence and competence (Figure 2)—two character traits central to trustworthiness impressions. I will show how different trusting behaviors will be generated by different weights on a partner's benevolence and competence, and how the model predicts that people should prefer a benevolent but incompetent partner over a malevolent one. The vulnerability model can be employed as is or as a utility function within more complex Bayesian frameworks for inferences in different social environments where actions are associated with subjective values and weighted by individual expectations of others' behaviors. This work provides an important building block for future theoretical and empirical work across a variety of research fields.

Keywords: trust, trustworthiness impressions, investment game, social learning, vulnerability

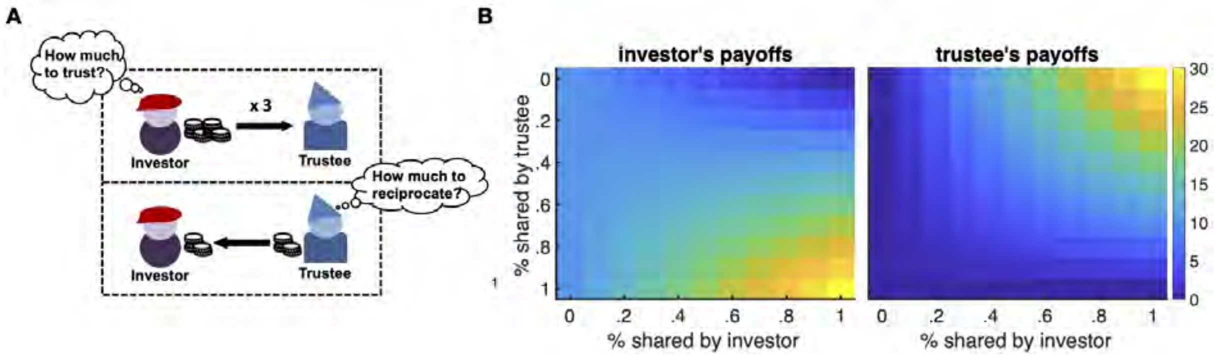


Figure 1: **Investment game and payoff structure.** The investment game (IG) is a sequential game where the investor receives an initial sum (endowment) and can decide whether to share any of this with the trustee (trust decision). The shared sum is then multipled (e.g., tripled) and passed on to the trustee who can decide to share back some of the received amount (reciprocity decision). In multiple one-shot or multi-round IGs, the same trial depicted is repeated multiple times with either different trustee or the same trustee, respectively (**A**). Investor's and trustee's payoffs at the end of each trial for each possible action (in % of the available sum). The payoff value grid clearly shows the dilemma of the investor as a function of economic interests (payoff values). While the investor is incentivized by the payoff structure to share higher amounts, the trustee is incentivized to keep more of the received amounts (**B**).

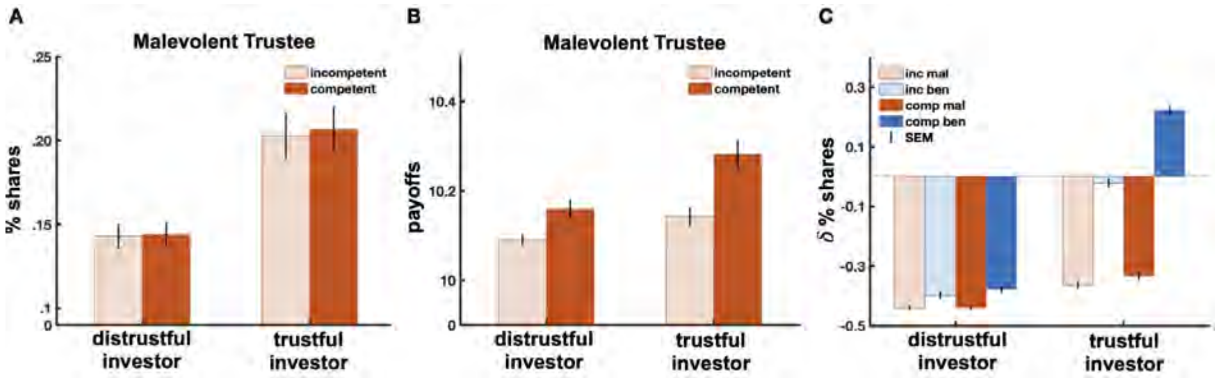


Figure 2: **Investment and payoffs comparison across trustee types.** Comparison of percentage of investments (A) and obtained payoffs (B) of distrustful (left) and trustful (right) investors for interactions with malevolent trustees who were additionally incompetent (light red) or competent (red). Differences of percentage shares between the default initial strategy of $i = .5$ and the first expectation-based investment i_t at $t = 3$ across all trustee types for distrustful (left) and trustful (right) investors. Inc: incompetent; Comp: competent; Mal: malevolent; Ben: benevolent.

Ambiguity Remains a Rare Skill for Learning about Others Even when Stimuli Carry Social Relevance

Asya Achimova & Marei Beukman
University of Tübingen
asya.achimova@uni-tuebingen.de

Ambiguity is an integral part of natural language and communication. Intuitively, ambiguity is an obstacle in successful comprehension, and speakers, driven by the Gricean maxims of co-operative communication (Grice, 1989), are expected to avoid it. Yet, psycholinguistic research shows that speakers do not avoid ambiguous descriptions even in situations where it seems to be communicatively appropriate (Ferreira, 2008; Wasow, 2015). Nonetheless, listeners often manage to successfully disambiguate the utterances, relying on context and world knowledge. Ambiguity may therefore be actually a sign of an efficient language system, where lightweight, well-formed pieces of language can be re-used to carry multiple meanings in order to optimize computational effort (Piantadosi et al., 2012).

Recent computational work focusing on the social aspects of ambiguity resolution have furthermore suggested that observing ambiguity resolution is indicative of particular beliefs a listener brings to a conversation (Achimova et al., 2022). These beliefs affect interpretation of utterance similarly to how beliefs and biases affect the perception of other stimuli, such as visual scenes (Boon & Davis, 1987). While participants in Achimova et al. (2022) successfully inferred the preferences of the listener upon observing her referential choice in a signaling game, only a subgroup of participants was able to use ambiguity for epistemic purposes, that is, for learning about others. Many others failed to successfully choose ambiguous utterances that allow learning about particular feature preferences of the listener upon observing their object choice amongst the alternatives. While the authors attribute the weak utterance choice performance in this task to limits of recursive reasoning, it is also possible that these results stem from the use of abstract stimuli.

In the current work, we amend the paradigm developed in Achimova et al. (2022) and use stimuli with faces rather than geometrical shapes to probe the ability to use ambiguity epistemically. We adopt experimental stimuli from Frank et al. (2016) and adjust them to match the original set up from Achimova et al. (2022). We hypothesize that the stimuli containing faces might promote a higher level of ambiguity use, similar to the effect of using socially relevant scenarios (Cosmides & Tooby, 1992) for the card selection task (Wason, 1968).

We report results from two online experiments: Experiment 1 (n = 100) targeted inference of preferences that made the listener choose a particular object, while Experiment 2 (n = 100) focused on utterance choice. Both experiments were carried out using the Prolific crowd-sourcing platform. Each experiment included two conditions: geometrical shapes and faces (Figure 1). Participants were randomly assigned to one of the conditions. They were not allowed to take part in both experiments.

The results, however, reveal that socially-relevant stimuli (faces) did not affect performance in either of the experiments. While participants still successfully inferred the simulated listener's preferences in Experiment 1, again only a subset of them managed to use ambiguity strategically to learn about the feature preferences of a simulated conversation partner (Figure 2). We conclude that epistemic ambiguity use remains a skill only available for some participants.

However, if ambiguity appears in conversation as a result of other factors, such as indirectness, all the participants might be able to infer the beliefs and preferences of others upon observing their ambiguity resolution behavior.

Keywords: referential ambiguity, social inference, RSA, signaling game, pragmatics

Figure 1: Both experiments contrasted stimuli composed of geometrical shapes vs. faces. Conditions varied between subjects.

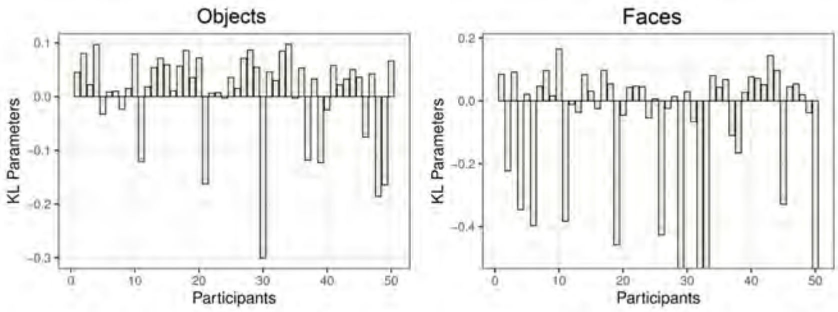


Figure 2: Positive values of the KL-value factor indicate epistemic use of ambiguous utterances. Negative values correspond to choosing primarily unambiguous utterances.

References

Achimova, A., Scontras, G., Stegemann-Philipps, C., Lohmann, J., & Butz, M. V. (2022). Learning about others: Modeling social inference through ambiguity resolution. *Cognition*, *218*, 104862.

Boon, J. C., & Davis, G. M. (1987). Rumours greatly exaggerated: Allport and postman's apocryphal study. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *19*(4), 430.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind: Evolutionary psychology and the generation of culture*, *163*, 163–228.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation: Advances in Research and Theory*, *49*, 209-246.

Frank, M., Goméz Emilsson, A., Peloquin, B., Goodman, N., & Potts, C. (2016). Rational speech act models of pragmatic reasoning in reference games. *PsyArXiv*.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*, 280-291.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly journal of experimental psychology*, *20*(3), 273–281.

Wasow, T. (2015). Ambiguity avoidance is overrated. In S. Winkler (Ed.), *Ambiguity: Language and communication* (p. 29-47). de Gruyter.

Gender associations elicited by masculine person-denoting nouns and indefinite pronouns

Hannah-Charlotte Bröder (Johannes Gutenberg University Mainz)

The discussion about the use of gender-fair language in German – a language with grammatical gender – is dominated by the debate about the so-called generic masculine (GM) (in German: "generisches Maskulinum"). This term refers to masculine person-denoting nouns (PNs) that are used a) when the gender of the persons referred to is not relevant, b) when mixed-gender groups are referred to, or c) for generalizations (Klann-Delius, 2005: 26). An example of the use of GMs is:

(1) **Bauarbeiter** [masc.] stehen früh auf. (**Construction workers** get up early.)

Numerous studies show that so-called GMs primarily make us think of men; therefore they impede the mental representation of non-male people (see among others De Backer & De Cuypere, 2012; Gygax et al., 2008; Irmen & Roßberg, 2004; Stahlberg & Szcesny, 2001). However, it remains an open question to what extent specific grammatical and semantic-lexical factors influence gender associations evoked by masculine PNs (Kotthoff & Nübling, 2018). To close this research gap, the project "Gender associations elicited by masculine person-denoting nouns and indefinite pronouns" (sub-project 2 of the DFG-project "Gender related practices in person reference: Discourse, grammar, cognition") plans several experiments to investigate the influence of said factors. In addition, indefinite pronouns like *jeder (everyone)* or *jemand (someone)* will also be examined, since they are also of masculine grammatical gender.

The focus of the presentation will be on the methods and results of the first two experiments. The first one involved a naming task (see among others Klein, 1988; Kusterle, 2011) to investigate to what extent particular grammatical and semantic-lexical factors influence gender associations evoked through PNs of masculine grammatical gender. The factors we tested were number (singular vs. plural), the stereotype of the PNs (female vs. neutral vs. male), and the PN type (occupation vs. role). Participants (n=44) had to name persons who had been introduced by a masculine PN in a preceding sentence, as outlined in the following example:

(2) **Friseure** [masc.] arbeiten meistens montags nicht. ___ und ___ machen jedoch eine Ausnahme […]. (**Hairdressers** usually do not work on Mondays. However, ___ and ___ make an exception [...].).

The advantage of the study at hand over previous studies is that the contexts were generic in a reference-semantic sense, i. e. the PNs did not refer to specific people (as they do in a sentence like: The hairdressers __ and __ […].).

In accordance with the hypotheses, PNs in the plural form, stereotypically female PNs and role PNs were most often associated with female persons. Apart from the interesting results of this first run of the experiment, the talk will also address an additional explorative experiment investigating gender associations evoked by indefinite pronouns. The results of both experiments will be discussed and interpreted.

Keywords: gender-fair language, generic masculine, gender associations, person-denoting nouns, indefinite pronouns.

References

De Backer, M., & De Cuypere, L. (2012). The interpretation of masculine personal nouns in German and Dutch: A comparative experimental study. *Language Sciences*, 34, 253–268.

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians and mechanics are all men. *Language and Cognitive Processes*, 23, 464–485.

Irmen, L., & Roßberg, N. (2004). Gender markedness of language. The impact of grammatical and nonlinguistic information on the mental representation of person information. *Journal of Language and Social Psychology*, 23(3), 272–307.

Klann-Delius, G. (2005). *Sprache und Geschlecht*. J.B. Metzler.

Klein, J. (1988). Benachteiligung der Frau im generischen Maskulinum – eine feministische Schimäre oder psycholinguistische Realität. In Oellers, N. (ed.), *Vorträge des Germanistentags Berlin 1987 (Bd. 1)* (pp. 310–319). Niemeyer.

Kotthoff, H., & Nübling, D. (2018). *Genderlinguistik. Eine Einführung in Sprache, Gespräch und Geschlecht*. Narr Francke Attempto.

Kusterle, L. (2011). *Die Macht von Sprachformen. Der Zusammenhang von Sprache, Denken und Genderwahrnehmung*. Brandes & Apsel.

Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, 52, 131–140.

Gender priming non-linguistic stimuli: the effect of word and sentence primes on the perception of male, female, and gender-mixed face pairs

Authors: Jonathan D. Kim[1], Anton Öttl[1], Pascal Gygax[2], Dawn M. Behne[1], Jukka Hyönä[3], Ute Gabriel[1]

[1]Department of Psychology, Norwegian University of Science and Technology, 7491 Trondheim, Norway
[2]Department of Psychology, University of Fribourg, Rue P.A. de Faucigny 2, 1700 Fribourg, Switzerland
[3]Division of Psychology, University of Turku, 20014 Turku, Finland

Research into gender representation in language comprehension has historically used linguistic, gender specific, primes and targets. As these results are not necessarily generalisable outside of the linguistic domain (Sato et al., 2016), more recent research has started to examine the effect of linguistic primes on non-linguistic targets. Further, the use of gender specific primes has left open the question as to whether mixed gender information is able to prime and/or to be primed. The current study aims to expand knowledge on both of these topics. Specifically, the Face Pair task (an approach to the Linguistic-Visual paradigm) was used across three experiments to examine the ability of mixed gender and gender specific information to both prime and be primed.

For the sake of clarity, research following the Linguistic-Visual paradigm uses linguistic stimuli to prime visual, but non-linguistic, targets. This approach has been found effective at determining the effects of gender priming. The Face Pair task itself is an innovative three-alternative forced choice task. Participants are presented with pairs of unambiguously gendered faces that are either both female, both male, or mixed gender. This task was specifically designed for use with priming tasks under the Linguistic-Visual paradigm.

The experiments undertaken in this study were run entirely in Norwegian as part of a wider project. A-priori power analysis for a pilot experiment indicated a sample size of N = 24 to find medium sized effects with alpha = 0.05, beta = 0.20/power = 0.80. The results of this pilot experiment were then used to guide selection for the other experiments in the project. We will discuss the results of three of these experiments, all of which use written (rather than aural) primes. Experiment 1 (N = 38) examined word primes (names and role nouns), while Experiments 2 (N = 45) and 3 (N = 32) examined a large range of one-sentence (Experiment 2) and two-sentence (Experiment 3) primes. In Experiment 1 both name and role noun primes were presented in female specific, male specific, and mixed/neutral gender forms. In Experiments 2 and 3, for reasons associated with the wider project, name primes were presented in female and male specific forms only while role noun primes were presented in gender neutral form only. For the purposes of this presentation, we will discuss the results for name and role noun primes for Experiment 1, but only name primes for Experiments 2 and 3. Data was analysed through general linear mixed effects regression and linear mixed effects regression using the lmer and glmer functions of the lme4 package (version 1.1-23; Bates et al., 2015).

For all experiments, and for both error rate and response time analyses, initial models were designed that captured the maximal fixed and random effects structure as justified by the design (i.e., all relevant experimental factors and their interactions, as well as random intercepts of relevant factors to each experiment, and random slopes of each experimental factor by each relevant random intercept). Further, mean responses to a within-block control prime (per participant and per face pair category) were included during model fitting to correct for by-participant motoric noise. Model refinement to find the model of best fit was done through forward-testing log-likelihood tests. In these tests all effects (both fixed and random) were added to an initially blank model, with their contribution to improving the model evaluated through reference to 1) a model excluding the effect, 2) competing models each testing the inclusion of a different effect, and 3) the initial model. The sequential inclusion of each relevant factor allowed for the model of best fit (i.e., the model that contained all the factors that were found to significantly explain variance) to be identified and then examined. Models that were overfitted (as indicated by having a singular fit) or which were unstable or failed to converge (therefore not producing reliable results) were excluded from consideration. To allow for discussions of priming effects prime gender, face pair gender, and their interaction were forced into final models when A) they would otherwise be excluded and B) their addition did not lead to overfitting, instability, or failure to converge. As such, the best model where these factors were naturally included are referred to as the models of best fit, while the best models where one or both factors had to be added are referred to as the final models.

The results of Experiment 1 (word primes; name pairs, e.g. *Nina og Cecilie* [Nina and Cecilie], and role nouns, e.g., *Sekretærene* [The secretaries]) indicated significant priming effects for Error Rate and Response Time. For Error Rate the results indicated that congruent primes led to less error than incongruent primes for all three face pair categories, with no differences between the name pair and role noun primes. For Response Time the results indicated slight differences between prime types. For name pair primes participants responded faster to congruent compared to incongruent gender-specific, but not mixed gender, faces. For role noun primes participants responded faster to congruent compared to incongruent male specific and mixed gender faces, but equally quickly regardless of prime gender to female faces.

The results of Experiment 2 (one-sentence primes; e.g. *Nina og Therese er uansvarlige* [Nina and Therese are irresponsible]) indicated no priming effects for Error Rate (p = .181) or Response Time (p = .081). Descriptively, participants produced slightly less errors for female faces following female compared to male primes, and responded slightly faster to male faces following male compared to female primes.

The results of Experiment 3 (two-sentence primes; e.g. *Therese og Ingrid er rike. De er på konserten* [Therese and Ingrid are rich. They are at the concert].) indicated a significant priming effect for Response Time (p = .010), with the results indicating that participants responded faster to congruent compared to incongruent primes for female (but not male) face pairs, and responded faster to mixed gender faces following male compared to female primes. No significant priming effect was found for Error Rate (p = .673). Descriptively, participants produced very low error rates overall, showing slightly lower error rates for female and male faces following congruent compared to incongruent primes.

Overall these results offer some support for the idea that gendered linguistic information can prime gendered non-linguistic targets. This is especially true for word primes, with the increasing amounts of linguistic non-gender information in the sentence primes moderating this effect in differing manners and to different degrees.

*Keywords***:** Gender primes, linguistic-visual paradigm, text-to-image priming, linguistic priming, face perception

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

R Core Team. (2018). *R: A Language and environment for statistical computing* (4.0.3) [Computer software]. https://www.R-project.org/

Sato, S., Gygax, P. M., & Gabriel, U. (2016). Gauging the impact of gender grammaticization in different languages: application of a linguistic-visual paradigm. *Frontiers in Psychology,* 140. https://doi.org/10.3389/fpsyg.2016.00140

The impact of social cognition on linguistic processing: Stereotypical information prevails over semantic encoding during reading.

Magali A. Mari[1], Misha-Laura Müller[1]
[1]Cognitive Science Center, University of Neuchâtel, Neuchâtel, Switzerland

The present study highlights the impact of gender and nationality stereotypes – i.e., socially shared beliefs about those specific categories – on language processing. Stereotypes function as heuristics as they automatically produce expectations about the members of a social category (Krieglmeyer & Sherman, 2012; Müller & Rothermund, 2014). As such, contents that are consistent with stereotypical expectations are processed more quickly than when contents violate a stereotype. Although classical research has well established that counter-stereotypical information affects language processing (see for e.g., Doherty & Conklin, 2017; Garnham et al., 2012; Oakhill et al., 2005; Pei et al., 2017; Sato et al., 2016), those past studies only focused on gender stereotypes. In the present study, we aimed at assessing the effects of other types of stereotypes, namely nationality-related stereotypes, while also offering a cross-cultural and cross-linguistic replication of previous finding with gender stereotypes.

Language processing is also affected by certain linguistic contents, such as definite descriptions (e.g., *the* scientist), which tend to elicit shorter processing times, while other linguistic contents, such as indefinite descriptions (e.g., *a* scientist) are less quickly processed (Frazier, 2006; Müller & Mari, 2021; Singh et al., 2016). In addition of being processed quickly, definite descriptions also entail a sense of familiarity and informational uniqueness (Heim, 1982; Schwarz, 2009; Roberts, 2002). Drawing from this perspective, if information about a social category is provided with definite description, the information might be considered as familiar or taken for granted. In the current study, we assessed whether manipulating linguistic inputs could affect the processing of counter-stereotypical information.

Two experiments tested separately the impact of gender and nationality stereotypes in the presence of definite and indefinite descriptions on information processing with a self-paced reading task. On the one hand, theoretical perspectives in social cognition propose that information violating a stereotype is costlier to process than information confirming a stereotype. On the other hand, theoretical accounts in psycholinguistic propose that definite descriptions facilitate information processing. By combining linguistic input and social information, we assessed whether one prevails over the other in language processing. If counter-stereotypical information elicits longer reading times, even when presented with definite descriptions, we could conclude that social expectancies have a stronger impact than linguistic input on information processing.

The stimuli consisted in two sentences written in French. The first sentence introduced a context and was then followed by a target sentence matching or violating a gender (Experiment 1) or a nationality-related (Experiment 2) stereotype. For example, in Experiment 1, participants read sentences like “Mary went to the garage last week. The/A mechanic$_{female/male}$ lent her a car”. Stimuli of nationality stereotypes included sentences like “Paul went to France/Japan last summer. The/A French/Japanese seemed very flirty”. The stimuli were pre-tested and selected from a list of 50 gender stereotypes and 90 nationality stereotypes. Forty subjects from the same population as our final sample indicated on a 5-point Likert scale how much they agreed with statements such as “Japanese people are flirty” or “Most mechanics are men”. Stereotypes that were close in average to the extreme points of the scale were used as stimuli in our study.

A total of 57 French-speaking Swiss adults (60.8% self-identified as woman, age mean: 23.87 years) from the region of Neuchâtel participated in Experiment 1. Overall, Experiment 1 showed that gender counter-stereotypical information yielded longer reading times than information confirming gender stereotypes. This replicates classical research on the effect of gender stereotypes on information processing with French-speaking Swiss participants.

Regarding the effect of definite descriptions, the results showed that reading times were faster only when information confirmed gender stereotypes. When counter-stereotypical information was introduced with definite descriptions (e.g., “the mechanic$_{male}$), participants’ reading times were significantly longer compared to the other conditions. This result shows that manipulating linguistic input failed to ease the processing of gender counter-stereotypes. On the contrary, definite descriptions combined with counter-stereotypes produced reading times that were significantly longer than any other conditions. This finding suggests that counter-stereotypical information was especially costly to process when introduced with definite descriptions. In parallel, presenting stereotypical information with definite description was not different from presenting information with indefinite description, suggesting that information confirming a stereotype is easily processed independently from the kind of linguistic input used.

Figure 1. *Reading times per conditions and types of stereotypes.*

(a) (b)

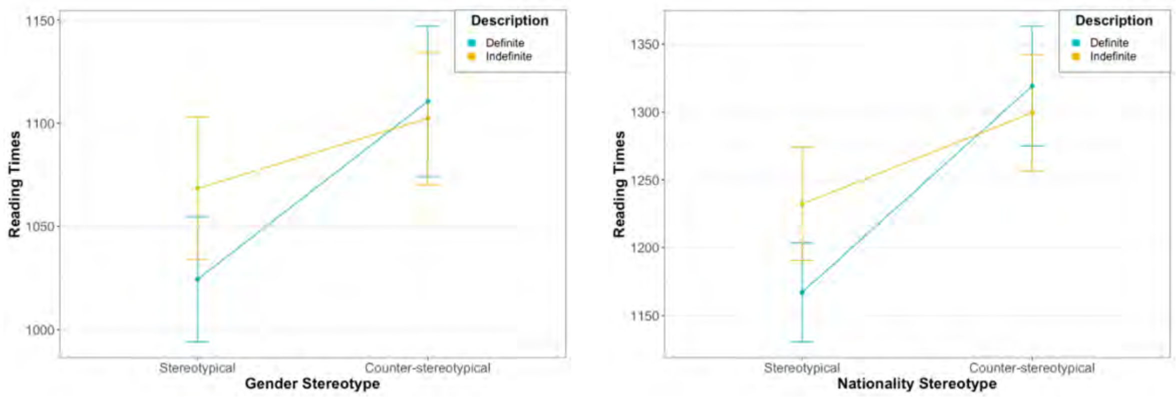


Figure 1. Mean reading times per conditions: stereotypical versus counter-stereotypical information introduced with definite versus indefinite descriptions. Error bars represent standard errors of the mean.

Experiment 2 included a total of 49 French-speaking Swiss adults (57.1% self-identified as woman, age mean: 23.06 years) from the region of Neuchâtel. When analysing the effect of stereotypes only, Experiment 2 revealed that reading times were longer for counter-stereotypical information compared to information confirming a stereotype. This result extends classical findings on gender stereotypes and provides evidence for a costly processing of other types of counter-stereotypes. Regarding the effect of definite descriptions, reading times of counter-stereotypes were not affected by the description: no difference was observed between reading times of counter-stereotypes introduced with definite or indefinite descriptions. However, when stereotypical information was introduced with a definite description, reading times were significantly faster than any other conditions. This finding suggests that stereotypical information was especially easy to process when introduced with definite descriptions.

Altogether, this study replicates and extends classical studies of the effect of stereotypes on information processing. The present findings suggest that counter-stereotypical information, related to gender or nationality stereotypes, is costly to process. In addition, while definite descriptions led to a fast processing of stereotypical information, they failed to facilitate the processing of counter-stereotypes. Together, these results provide further evidence in favour of a view that language comprehension interacts with other modules of cognition. Our findings reveal that social expectancies have a stronger impact than linguistic input on information processing. In conclusion, the current study provides cross-disciplinary perspectives for the investigation of the interaction between social and linguistic modules of cognition.

References

Frazier, L. (2006, October). The big fish in a small pond: Accommodation and the processing of novel definites. In *Ohio State University Presupposition Accommodation Workshop*, Columbus, Ohio.

Doherty, A., & Conklin, K. (2017). How gender-expectancy affects the processing of “them”. *The Quarterly Journal of Experimental Psychology, 70*(4), 718-735. https://doi.org/10.1080/17470218.2016.1154582

Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P., & Oakhill, J. (2012). Gender representation in different languages and grammatical marking on pronouns: When beauticians, musicians, and mechanics remain men. *Discourse Processes, 49*(6), 481–500. https://doi.org/10.1080/0163853X.2012.688184

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD dissertation. Amherst: GLSA, University of Massachussetts.

Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *Journal of Personality and Social Psychology*, *103*(2), 205-224. https://doi.org/10.1037/a0028764

Müller, M.-L., & Mari, M. A. (2021). Definite descriptions in the light of the comprehension vs. acceptance distinction: Comparing self-paced reading with eye-tracking measures. *Frontiers in Communication, 6*(634362), 1-17. https://doi.org/10.3389/fcomm.2021.634362

Müller, F., & Rothermund, K. (2014). What does it take to activate stereotypes? Simple primes don’t seem enough: A replication of stereotype activation (Banaji & Hardin, 1996; Blair & Banaji, 1996). *Social Psychology, 45*(3), 187-193. https://doi.org/10.1027/1864-9335/a000183

Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition, 33*(6), 972–983. https://doi.org/10.3758/BF03193206

Sato, S., Gygax, P. M., & Gabriel, U. (2016). Gauging the impact of gender grammaticization in different languages: Application of a linguistic-visual paradigm. *Frontiers in Psychology, 7*(140). https://doi.org/10.3389/fpsyg.2016.00140

Schwarz, F. (2009). *Two Types of Definites in Natural Language*. *Open Access Dissertations*, 122. University of Massachusetts Amherst. https://scholarworks.umass.edu/open_access_dissertations/122/?utm_source=scholarworks.umass.edu%2Fopen_access_dissertations%2F122&utm_medium=PDF&utm_campaign=PDFCoverPages

Singh, R., Fedorenko, E., Mahowald, K., & Gibson, E. (2016). Accommodating presuppositions is inappropriate in implausible contexts. *Cognitive Science, 40*(3), 607–634. https://doi.org/10.1111/cogs.12260

Wang, P., Yang, Y.P., Tan, C.H., Chen, Q.W., & Cantfort, T. (2017). Gender Stereotype activation versus lexical semantic activation: An ERP study. *The Journal of General Psychology, 144*(4), 283-308. https://doi.org/10.1080/00221309.2017.1310714

Session 7

Reproducing Children's Category Exemplars for Fruit Categories

Pablo Léon-Villagrá[1] **Isaac Ehrlich**[2]
Christopher G. Lucas[3] **Daphna Buchsbaum**[1]

[1]Brown University, USA
[2]University of Toronto, Canada
[3]University of Edinburgh, UK

Worked Examples as Application of Analogical Reasoning in Intelligent Tutoring and their Effects on SQL Competencies

Thaler, Anna Magdalena[1]; Mitrovic, Antonija[2]; Schmid, Ute[1]

[1] Cognitive Systems Group, University of Bamberg, Germany

[2] Department of Computer Science and Software Engineering, University of Canterbury, New Zealand

The acquisition of problem-solving skills such as solving mathematical equations or programming demand a substantial number of cognitive schemata. Problem-based learning has been found to improve general problem-solving strategies and critical thinking and outperform lecture-based environments (Yew & Goh, 2016). Presenting worked examples (WEs) facilitates the acquisition of required knowledge (Sweller & Cooper, 1985). WEs display all required steps to solve the problem and this insight can be transferred to the current problem (VanLehn, 1998). The mapping process can be viewed as analogical reasoning and requires isomorphic Wes with the same underlying structure as the current problem (Gick & Holyoak, 1983). Learning from analogies allows for building problem-solving schemata and increases knowledge by schema abstraction (Gentner & Maravilla, 2018). Analogical comparison is assumed to be more effective for far transfer problem solving than near transfer compared to methods such as self-explaining (Nokes-Malach et al., 2013). An approach for an automatic generation of isomorphic examples was proposed by Zeller & Schmid (2016). There, the generated isomorphic example was constrained such that it highlights identified misconceptions. In the domain of programming, examples are especially effective to support the acquisition of more general program schemes (Pirolli & Anderson, 1985).

Najar et al. (2015) incorporated WEs in an existing intelligent tutoring system (ITS) called SQL-Tutor (Mitrovic & Ohlsson, 1999), and found that novices investigated only Wes, while advanced students also focused on additional information such as database schema. SQL-Tutor employs Constraint-Based Modeling, an approach where the domain knowledge is formulated as a set of constraints, which helps identify incorrect or incomplete knowledge if the student violates these constraints (Mitrovic & Ohlsson, 1999). The system allows individual feedback on several levels of increasing degree of assistance. To ensure structural similarity independent of the complexity of posed SQL problems, WEs are created according to problem templates with identical underlying structures but different surface structures (Mathews & Mitrovic, 2007). The use and effectiveness of WEs in ITSs depend on the learner's prior knowledge (Najar et al., 2015). Additionally, it is postulated that motivation and metacognitive strategies play a role in the proper use of feedback within an ITS (Hull & Du Boulay, 2015). Providing examples can enhance students' motivation to interact with the problem domain more thoroughly (Narciss & Huth, 2006). High self-efficacy expectations, i.e. the perceived competence in performing a task, can determine further self-regulatory learning strategies and seem to strongly mediate the use of provided materials (Hull & Du Boulay, 2015). Apart from observable performance measures, ways to increase positively correlated subjective measures such as self-efficacy for programming tasks should be explored further.

We are currently conducting an experiment (July 2022) to investigate the effects of WEs as an implementation of analogical reasoning on the perceived and measurable competency in a programming task. Students will be formulating SQL queries using SQL-Tutor (Mitrovic, 2003). We use pre/post-tests to determine students' conceptual and procedural knowledge of SQL, as well as self-efficacy expectations towards SQL skills before/after the study session. There are three groups of students. In the case of submitting an erroneous solution, the experimental group 1 (EG1) students receive an isomorphic WE from the same database, while the experimental group 2 (EG2) students receive an isomorphic WE from a different database. This distinction between databases aims to test previous findings on the hypothesis that analogical comparison works better on far transfer problems, so we expect higher scores for EG2. The control group students receive the full solution instead of a WE. After receiving WEs or the full solution, the students will have one more attempt to solve the problem or copy the correct response accordingly. We hypothesize that the analogical mapping of insights of the WEs to the target problem encourages a deeper cognitive analysis and should have positive effects on performance. We also hypothesize that WEs shift the focus from pointing out students' single errors to viewing the ITS support as an additional learning opportunity, and so WEs should have a positive effect on students' self-efficacy expectations.

Keywords: Intelligent Tutoring Systems, SQL-Tutor, Analogical Reasoning, Worked Examples

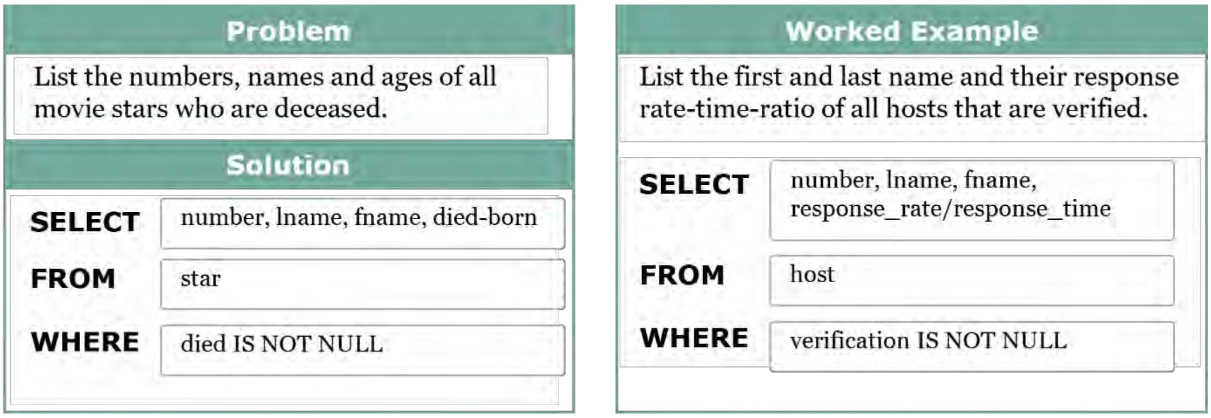


Figure 1. Example of SQL problem with solution (left) and WE from a different database (right).

References

Gentner, D., & Maravilla, F. (2018). Analogical reasoning. In L. J. Ball & V. A. Thompson (Eds.), *International Handbook of Thinking & Reasoning* (pp. 186–203).

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Hull, A., & Du Boulay, B. (2015). Motivational and metacognitive feedback in SQL-Tutor *Computer Science Education*, *25*(2), 238–256.

Lishinski, A., Yadav, A., Good, J., & Enbody, R. (2016). Learning to Program: Gender Differences and Interactive Effects of Students' Motivation, Goals, and Self-Efficacy on Performance. In J. Sheard (Ed.), *Proc. of the 2016 ACM Conf. on International Computing Education Research* (pp. 211–220). Association for Computing Machinery.

Mathews, M., & Mitrovic, A. (2007). The Effect of Problem Templates on Learning in Intelligent Tutoring Systems. *0922-6389*.

Mitrovic, A. and Ohlsson, S. (1999) Evaluation of a Constraint-Based Tutor for a Database Language. *Artificial Intelligence in Education*, 10(3-4), 238-256.

Najar, A. S., Mitrovic, A., & Neshatian, K. (2015). Eye Tracking and Studying Examples: How Novices and Advanced Learners Study SQL Examples. In *Journal of Computing and Information Technology* (pp. 171–190).

Narciss, S., & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction*, *16*(4), 310–322.

Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, 28(4), 1237-1263.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. Canadian Journal of Psychology, 39(2), 240-272.

Sweller, J., & Cooper, G. A. (1985). The Use of Worked Examples as a Substitute for Problem Solving in Learning Algebra. In *Cognition and Instruction* (Vol. 2, pp. 59–89).

VanLehn, K. (1998). Analogy Events: How Examples are Used During Problem Solving. *Cognitive Science*, *22*(3), 347–388.

Yew, E. H., & Goh, K. (2016). Problem-Based Learning: An Overview of its Process and Impact on Learning. Health Professions Education, 2(2), 75–79.

Zeller, C. & Schmid, U. (2016). Automatic Generation of Analogous Problems to Help Resolving Misconceptions in an Intelligent Tutor System for Written Subtraction. Workshops Proc. 24th Int. Conference on Case-Based Reasoning (ICCBR 2016).

Effects of inclusive dance training on body representations and body-related concepts in dancers with different bodily conditions

Bettina Bläsing, Technical University Dortmund, Faculty of Rehabilitation Sciences, Music and Movement in Rehabilitation and Education; in collaboration with Gerda König & Gitta Roser (Din A13, M.A.D.E.)

Based on perspectives that acknowledge the body as fundamental for cognitive functions, has often been claimed that movement-based, information-rich domains such as sports and the performing arts represent valuable fields for investigating human cognition in the real world (Beilock, 2008; Bläsing et al., 2012; Yarrow et al., 2009). An aspect that has not yet been considered by many authors is the situation of performers with disabilities. whose individual forms of double expertise is specific beyond sophisticated performance in the given domain, as it also includes the mastery of challenges provided by their specific bodily or sensory condition (Quinten et al., 2022; Whatley, 2007). Even though the relevance of the body is indisputable in embodied cognition, the degree to which the physical body or its neurocognitive representations are granted explanatory value is under discussion (Alsmith & De Vignemont, 2012). Different types of body representation have been described with regards to their origins and functions (De Vignemont, 2010; Longo, 2015), with most consistent evidence supporting a differentiation of body schema and body image. Haggard and Wolpert (2005) ascribe seven main characteristics to body schema, claiming that it is spatially coded, modular, continuously updated with movement, adaptable to changes in the properties of the physical body, supra-modal, coherent over time, and interpersonal (serving as model for understanding others' bodies and actions). Evidence exists that in persons with missing limbs, functional adaptations in action execution can modify the modular structure of body schema, whereas phantom limbs might maintain the primary modularity, even in persons who have never experienced a complete normative body (Bläsing et al., 2010).

The study presented here has been conducted in collaboration with the dance company DinA13 and M.A.D.E. (Mixed-Abled Dance Education), a professional training course for dancers with non-/normative bodies and with/without physical disabilities. The project focuses on body representations and body- / movement-related concepts in the context of inclusive dance. Two data collections were carried out, at the beginning of the course and approximately 30 months later. The study aims to explore in what ways body representations and underlying body- and action-related concepts reflect participants' physical conditions, and to what extent the shared experience of inclusive dance training affects participants' body representations and concepts. Three methodological approaches were combined into a triangular (quantitative / qualitative) study design. In task 1, structural relationships between predefined body- and movement-based concepts were analyzed using a quantitative paradigm that has previously used to investigate movement structures in ling term memory (e.g., Bläsing et al., 2009; Land et al., 2013). In this task, 22 verbal labels for body parts, physical activities (e.g. rolling, running, pulling) and more abstract body-based concepts (e.g. focus, axis, direction) had to be sorted (Bläsing et al., 2010). The results of the statistical evaluation (cluster analysis) can be interpreted as representing knowledge structures in participants' long-term memory (Land et al., 2013). Individual cluster solutions were found to reflect participants' bodily conditions and motor experience (e.g., use of assistive technologies) as well as their training experience in (inclusive) dance.

Two qualitative tasks were added to gain a deeper understanding of the body- and movement-related concepts the participants had acquired through their individual bodily conditions and training experiences. In the first task, verbal terms (corresponding to those used in task 1) were presented together with a sketch of a human body figure (Majid & Van Staden 2015) on an otherwise blank page. Participants were asked to sketch, draw or write whatever they associated with the term, using the human figure as template. In the third task, only the verbal term was written on a blank page, and participants were again asked to illustrate their associations by sketching, drawing and writing. Combined findings from the three tasks reveal influences of inclusive dance experience on individual participants' body representation structures (body schema), as well as body-based (body parts, movement types) and more abstract concepts (e.g., perception, aesthetics).

*Keywords***:** body schema, action concepts, dance, disability, expertise

References

Alsmith, A. J. T., & De Vignemont, F. (2012). Embodying the mind and representing the body. *Review of Philosophy and Psychology*, *3*(1), 1-13.

Beilock, S. L. (2008). Beyond the playing field: Sport psychology meets embodied cognition. *International Review of Sport and Exercise Psychology*, *1*(1), 19-30.

Bläsing, B., Calvo-Merino, B., Cross, E., Jola, C., Honisch, J., & Stevens, K. (2012a). Neurocognitive control in dance perception and performance. *Acta Psychologica, 139*(2), 300-308.

Bläsing, B., Schack, T., & Brugger, P. (2010). The functional architecture of the human body: assessing body representation by sorting body parts and activities. *Experimental Brain Research, 203*(1), 119-129.

Bläsing, B., Tenenbaum, G., & Schack, T. (2009). The cognitive structure of movements in classical dance. *Psychology of Sport and Exercise*, *10*(3), 350-360.

De Vignemont, F. (2010). Body schema and body image – pros and cons. *Neuropsychologia 48*, 669-680.

Haggard, P., & Wolpert, D. M. (2005). Disorders of body scheme. In In Freund, H.J., Jeannerod, M., Hallett, M., Leiguarda R.,(Eds.), *Higher-Order Motor Disorders.* Oxford University Press, Oxford, pp. 261–272.

Land, W., Volchenkov, D., Bläsing, B. E., & Schack, T. (2013). From action representation to action execution: exploring the links between cognitive and biomechanical levels of motor control. *Frontiers in Computational Neuroscience*, *7*, 127.

Longo, M. R. (2015). Types of body representation. In Y. Coello, & M. H. Fischer (Eds.), *Perceptual and Emotional Embodiment* (pp. 125-142). Routledge.

Majid, A., & Van Staden, M. (2015). Can nomenclature for the body be explained by embodiment theories?. *Topics in cognitive science, 7*(4), 570-594.

Quinten, S., Bläsing, B., & Whatley, S. (2022). Dance and Disability. *Frontiers in Psychology, 1075.*

Whatley, S. (2007). Dance and disability: The dancer, the viewer and the presumption of difference. *Research in Dance Education*, *8*(1), 5-25.

Yarrow, K., Brown, P., & Krakauer, J. W. (2009). Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nature Reviews Neuroscience, 10*(8), 585–596.

Open(ing) Knowledge – Concepts and Discourses regarding Open Science

René Dutschke[1], Alexander Lasch[1], Simon Meier-Vieracker[1], Stefan Scherbaum[1], Sophia Seemann[1], Lucie Weigelt[1]

Central institutions of the so-called knowledge society, such as universities, libraries, and schools, are heavily affected by growing digitization. Alongside the technological advancement, individual stakeholders' attitudes and beliefs play a crucial role in this process. To account for that, a psychological and a linguistic perspective are combined to look upon individual construct-systems (Kelly, 1955) as well as discursive representation (Bubenhofer, 2009) of the disruptive and innovative potentials of digitization.

Scientific research and education are changing under the influence of digitization and the increasing adaptation of "Open Science"-practices (Meier et al., 2020). Qualitative and quantitative linguistic methods are combined to identify core concepts in Open Science discourses. A mixed-methods approach is applied to study individual beliefs about those concepts and their affective connotations. Data from Repertory Grid interviews (Kelly, 1955) and a triad comparison test is used to construct cognitive-affective maps (Thagard, 2012).

In semi-structured Repertory Grid interviews members of the academic community, including proponents and sceptics of the Open Science movement, are questioned regarding their personal views on good scientific practice. By sequentially comparing triads of people participants identified as good or poor researchers, constructs are elicited, which are, in a broader sense, dichotomous reference axes individuals use to make sense of the world (Kelly, 1955). Since the goal of this method is to capture the individual belief system, no objective definition of "good" or "poor" research is given. Rather it is left to the subjective judgement of participants. The structure of participants construct system can be studied qualitatively and quantitatively (e.g. Burr et al., 2020; Höft et al., 2019) and will be visualized as a cognitive-affective map (Thagard, 2012, 2015; Homer-Dixon et al., 2014)

In a computerized triad test participants are asked to categorize concepts identified as central in a corpus of Open Science texts regarding their similarity. In each triad of words two words must be chosen which are similar to each other and different then the third. In addition, the concepts are rated regarding their emotional valence. The calculated similarity of concepts along with their emotional ratings are subsequently used to construct maps to visualize the cognitive-affective structure regarding the concepts. In a first study, retest-reliability as well as external validity of the triad test is estimated.

In the course of the conference talk an outline of the research project with specific focus on its interdisciplinary scope and mix of qualitative and quantitative methods will be presented. First findings from the triad test validation study will be used to illustrate a tool to visualize belief systems.

Keywords: Cognitive-affective maps, Repertory Grid, Open Science, corpus linguistics

[1] Technische Universität Dresden; authors listed in alphabetical order

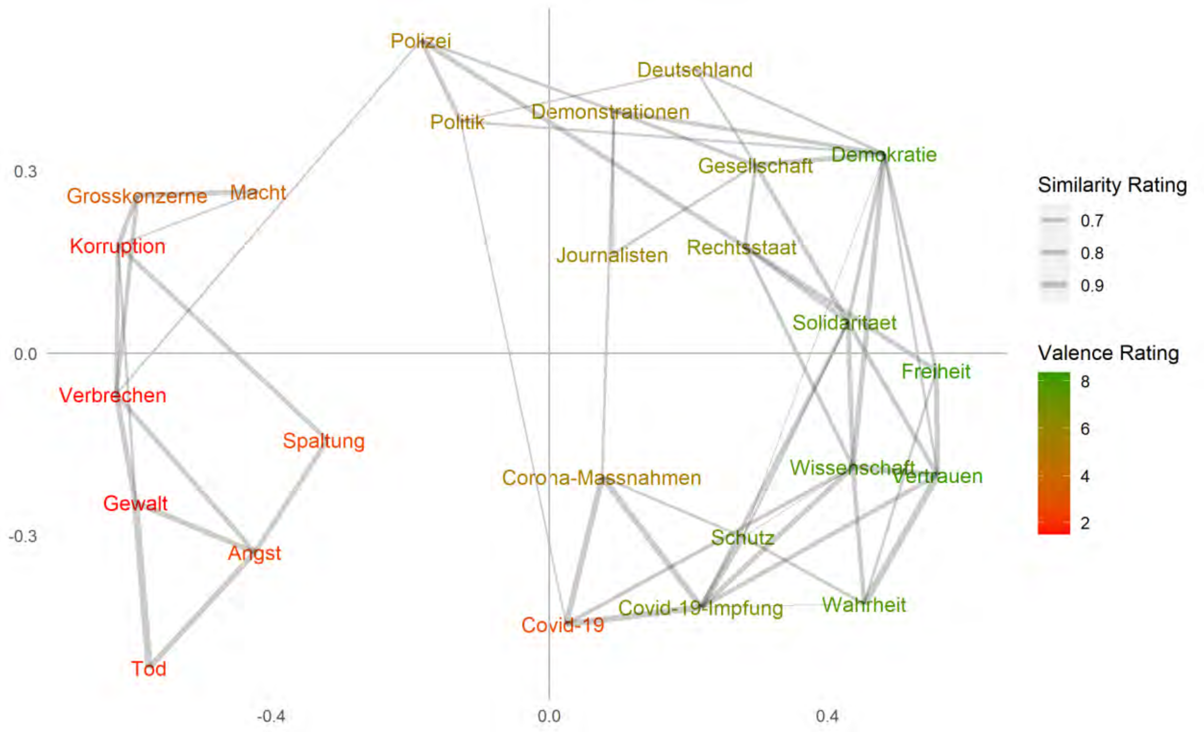


Figure 1. Example of visualization of cognitive-affective map from a pilot study (on beliefs about the Corona pandemic)

References

Bubenhofer, N. (2009). *Sprachgebrauchsmuster: Korpuslinguistik als Methode der Diskurs- und Kulturanalyse* (Vol. 4). Walter de Gruyter. https://www.degruyter.com/document/doi/10.1515/9783110215854/html

Burr, V., King, N., & Heckmann, M. (2020). The qualitative analysis of repertory grid data: Interpretive Clustering. *Qualitative Research in Psychology*. https://doi.org/10.1080/14780887.2020.1794088

Höft, N., Heckmann, M., & Jankowicz, D. (2019). Systematic Integration of Quantitative Measures into the Qualitative Content Analysis of Constructs. *Journal of Constructivist Psychology*, *32*(4), 345–369. https://doi.org/10.1080/10720537.2018.1515044

Homer-Dixon, T., Milkoreit, M., Mock, S. J., Schröder, T., & Thagard, P. (2014). The Conceptual Structure of Social Disputes: Cognitive-Affective Maps as a Tool for Conflict Analysis and Resolution. *SAGE Open*, *4*(1). https://doi.org/10.1177/2158244014526210

Kelly, G. A. (1955). *Personal construct psychology*. Norton.

Meier, S., Viehhauser, G., & Sahle, P. (Eds.). (2020). *Rekontextualisierung als Forschungsparadigma des Digitalen* (Vol. 14). BoD – Books on Demand.

Thagard, P. (2012). Mapping minds across cultures. In R. Sun (Ed.), *Grounding social sciences in cognitive sciences* (pp. 35–62). Cambridge, MA: MIT Press.

Thagard, P. (2015). The Cognitive–Affective Structure of Political Ideologies. In B. Martinovsky (Ed.), *Emotion in Group Decision and Negotiation* (Vol. 7, pp. 51–71). Springer Netherlands. https://doi.org/10.1007/978-94-017-9963-8_3

Tutorials

Tutorial Proposal: Neural Dynamics For Embodied Cognition

Abstract

Both in an evolutionary and a developmental perspective, cognition emerges from sensorimotor behaviors that become increasingly invariant and abstract. These sensorimotor origins of cognition explain why mental processes are so intimately intertwined with perceptual and motor processes, and share key properties, including, most prominently, the continuity of processing in state and time, and the dynamic stability of functionally meaningful states.

The mathematical framework of Dynamic Field Theory (Schöner, Spencer, and the DFT research group, 2016) makes these hypotheses explicit by postulating that cognition is based on the activation dynamics of neural populations organized as strongly recurrent neural networks that stabilize neural representations. Instabilities of such neural representations generate the state transitions that build sequences of mental and motor acts.

DFT enables models of cognition in two flavors. (1) Psychophysical experiments can be accounted for in neural process models that instantiate specific cognitive and behavioral competences such as visual attention, working memory, change detection, executive control, and many more. Current research probes how far toward higher cognition such embodied neural process accounts may reach (Richter, Lins, Schöner, 2021). (2) DFT models can also be used to generate behavior and thought in autonomous agents that are situated in structured environment and ultimately produce motor behavior (Tekülve et al., 2019).

Learning the concepts of DFT may enable cognitive scientists to broaden the reach of their theoretical work toward neurally grounded models. Integration across many different types of processes from the sensorimotor level to higher cognition is a key element of DFT. This may provide a useful complement to students typical training in computational concepts of system integration. Ultimately, understanding the need for stability in neural approaches to cognition will be a unique insight from this short workshop.

References

Schöner, G., Spencer, J. P., and the DFT Research Group (2016): *Dynamic Thinking: A Primer on Dynamic Field Theory*. Oxford University Press.

Tekülve, J., Fois, A., Sandamirskaya, Y., Schöner, G. (2019): Autonomous sequence generation for a neural dynamic robot: Scene perception, serial order, and object-oriented movement. *Frontiers in Neurorobotics*, 13, 1-27.

Richter, M., Lins, J., Schöner, G. (2021). A neural dynamic model for the perceptual grounding of spatial and movement relations. *Cognitive Science*, 45, e13405.

Event Format

The tutorial will comprise a lecture (one hour) that provides the conceptual foundations and connects to participant's background knowledge, followed by one hour of interactive hands-on exercise in which participants can be begin to build their own neural-dynamic architectures using the open-source simulation tool *cedar*.[1] Participants will work in teams of two on their own laptops and will receive support by a team of assistants coordinated by Dr. Tekülve.

Organizers

Prof. Dr. Gregor Schöner	Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany
Dr. Jan Tekülve	Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany
Dr. Mathis Richter	Neuromorphic Computing Lab, Intel, Munich, Germany

Gregor Schöner is the director of the Institute for Neuroinformatik (INI) at the Ruhr-Universität Bochum, where he holds the chair *Theory of Cognitive Systems*. In his long interdisciplinary careers he has contributed to theories of motor control and embodied cogntion. He is one of the originators of DFT.

Jan Tekülve is a postdoctural researcher at the INI where he directs a number of students on DFT projects. His earlier dissertation work developed a DFT account of intentionality. He is one of the developers of the simulation system, *cedar*, utilized in this tutorial.

Mathis Richter is a postdoctoral researcher at the Neuromorphic Computing Lab, Intel. He is building on DFT ideas to develop applications of neuromorphic hardware in the area of autonomous robotics.

[1] *Cedar* is available for all major operating systems at `cedar.ini.rub.de`.

Title: Meta-Learned Models of Cognition

Summary:
Research in cognitive psychology and neuroscience relies on computational models to study, analyze and understand human learning. Traditionally, such computational models have been hand-designed by expert researchers. In a cognitive architecture, for example, researchers provide a fixed set of structures and a definition of how these structures interact with each other. In a Bayesian model, researchers instead specify a prior and a likelihood function, which in combination with Bayes' rule, fully determine the model's behavior. The framework of meta-learning offers a radically different approach for constructing computational models of learning. In this framework, learning algorithms are acquired – i.e., they are themselves learned – through repeated interactions with an environment instead of being a priori defined by a researcher.

Recently, psychologists have started to apply meta-learning in order to study human learning. In this context, it has been demonstrated that meta-learned models can capture a wide range of empirically observed phenomena that could not be explained otherwise. They, amongst others, reproduce human biases in probabilistic reasoning [1], discover heuristic decision-making strategies used by people [2], and generalize compositionally on complex language tasks in a human-like manner [3]. The goal of this tutorial is to introduce the general ideas behind meta-learning, highlight its close connections to the Bayesian inference, and discuss its advantages and disadvantages relative to other modeling frameworks. We will furthermore cover how to implement these models using the PyTorch framework.

[1] Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological review*, *127*(3), 412.

[2] Binz, M., Gershman, S. J., Schulz, E., & Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological review*.

[3] Lake, B. M. (2019). Compositional generalization through meta sequence-to-sequence learning. *Advances in neural information processing systems*, *32*.

Information about the event's format:
The tutorial will take the form of an interactive lecture that is divided into two parts. In the first part, I will introduce the general concepts and ideas of meta-learning and explain how it can be used to construct models of human cognition. In the second part, I will show how to implement such models in PyTorch by walking through a code example.

Names and affiliations of the organizers:
Marcel Binz, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

I have written a doctoral thesis on how to use meta-learning to construct cognitive models, published multiple papers in this area, and have several others under review or in preparation.

Poster Session

Signal enhancement or noise exclusion?
Effect of temporal preparation on perceptual processes

Janina Balke, Bettina Rolke, & Verena C. Seibold
University of Tübingen, Germany

When a warning signal is presented before an upcoming event, it induces a process of temporal preparation, that is, the anticipation of and preparation for a future event. Previous studies have established that temporal preparation facilitates perceptual processing in perceptually demanding tasks (e.g., Rolke, 2008). Based on the Framework of the *Perceptual Template Model* (Lu & Dosher, 2005), we investigated whether this perceptual effect of temporal preparation is caused by (one of) two selective attention mechanisms: First, an enhancement of the strength of the task-relevant input (i.e., the target; *signal enhancement*) and second, a reduction or suppression from interfering input (i.e., a distractor; *external noise exclusion*). In this experiment, we combined manipulations of the assumed attentional mechanisms (see Lu & Dosher, 2005) with a manipulation of temporal preparation.

Participants' (N = 8) task was to report the orientation of a Gabor patch (tilted left or right) which was presented in between visual noise masks (see Fig. 1). To measure stimulus enhancement and external noise exclusion, we varied the signal contrast of the Gabor patch as well as the percentage of noise (i.e., the proportion of white/black pixels) in the mask, respectively. To vary the degree of temporal preparation, we implemented a so-called *constant foreperiod paradigm*. There, a warning signal is presented before the actual stimulus, which participants can use to prepare for the upcoming target. To avoid priming effects on the mask and the Gabor patch, we used an auditory warning signal in the current study. The interval between warning signal and stimulus – the foreperiod (FP) – can be either short or long and is kept constant within blocks of trials. This typically leads to faster reaction time (RT) and higher accuracy in the short compared to the long FP condition, indicating better temporal preparation in the former case (e.g., Rolke, 2008). We calculated the threshold of signal contrast that was necessary to obtain 75% correct responses for all noise conditions, and we compared the thresholds between the short and the long FP condition. A modulation of the lower or the higher part of the signal contrast function allows differentiating between an effect of temporal preparation on signal enhancement or noise exclusion, respectively (Lu & Dosher, 2005).

A repeated-measures ANOVA with the factors *percentage of external noise* and *FP* on the threshold of signal contrast revealed that the signal contrast increased systematically with the percentage of external noise, $F(6, 42) = 21.37$, $p < .001$, $\eta_p^2 = .75$ (see Fig. 2). Importantly, the required signal contrast to obtain 75% correct responses was higher for the short FP (M = 2.31, SD = 1.14) than for the long FP (M = 2.14, SD = .93), $F(1, 7) = 5.81$, $p = .047$, $\eta_p^2 = .45$. There was no interaction between the two factors, $F(6, 42) = 1.23$, $p = .31$. Additional repeated-measures ANOVAs for RT and correct responses revealed that RT tended to be faster, $F(1, 7) = 5.15$, $p = .058$, and accuracy tended to be lower, $F(1, 7) = 4.31$, $p = .076$, in the short FP condition in comparison to the long FP condition.

Even though our results do not allow distinguishing between a FP effect on signal enhancement and noise exclusion, they provide new insights into the mechanism of temporal preparation. In particular, the observation that a higher signal contrast was required in the short rather than the long FP condition contradicts previous studies (e.g., Rolke, 2008) that revealed higher accuracy in the short compared to the long FP condition. Importantly, in these previous studies backward masking was used whereas in the present study both forward and backward masking was used. If temporal preparation unselectively influenced the first presented stimulus, it would either increase the strength of the noise (forward) mask, or of the signal. If the mask is presented first (as in our study), this would lead to a larger required signal contrast to overcome the noise; if, however, the signal is presented first, this would lead to a stronger signal and thus higher accuracy. Thereby, we can explain why temporal preparation leads to higher accuracy in backward masking and the adverse effect in the present study. Follow-up experiments with systematic variations of the type of masking could help to answer the question whether temporal preparation influences perception via an unselective improvement of the first stimulus or via selective mechanisms like signal enhancement or noise exclusion.

Keywords: attention, temporal preparation, psychophysics

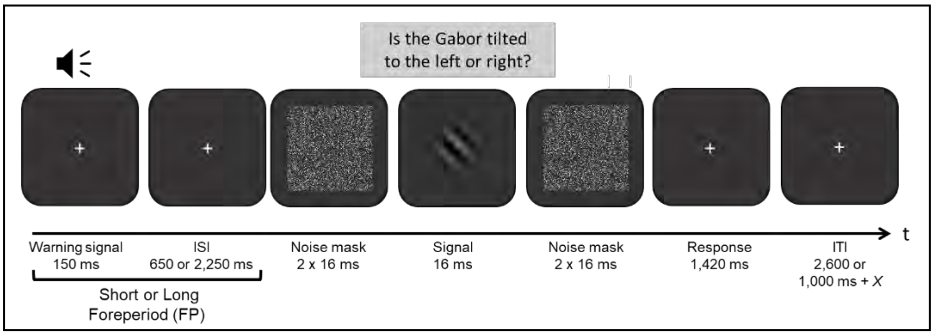


Figure 1. Illustration of an experimental trial (procedure adapted from Lu & Dosher, 2005). Participants' task was to report the orientation of the Gabor patch (here denoted as 'signal'). We manipulated temporal preparation via two constant foreperiods (short - 800 ms / long - 2,400 ms), the percentage of noise in the mask, and the signal contrast in the Gabor patch.

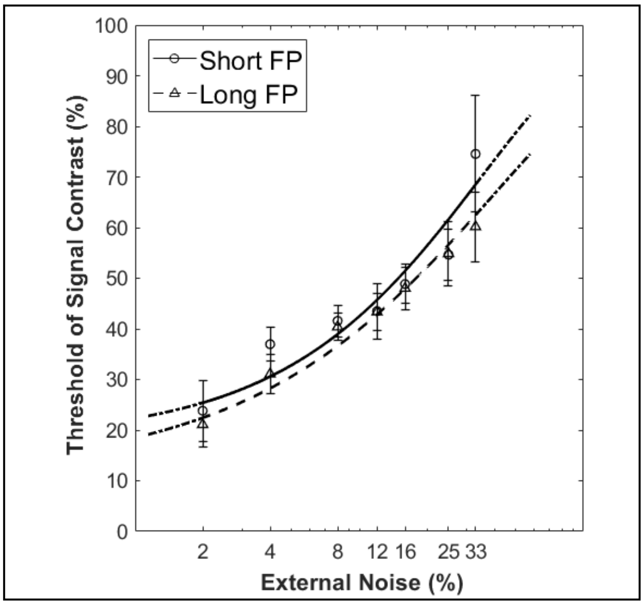


Figure 2. Results for the threshold of signal contrast (i.e., the amount of signal contrast that is required to obtain 75 % correct responses) as a function of the percentage of external noise in the mask, separately for the short and the long foreperiod (FP) condition (circles and triangles, respectively). The thresholds for each FP condition were fitted to psychometric functions (solid and dotted line), using the *psignifit* toolbox version 2.5.6 for Matlab, which implements the maximum-likelihood method described by Wichmann and Hill (2001).

References

Lu, Z. L., & Dosher, B. A. (2005). External noise distinguishes mechanisms of attention. In L., Itti, G., Rees, G., & J. K. Tsotsos, (Eds.), *Neurobiology of attention* (pp. 448-453). Academic Press.

Rolke, B. (2008). Temporal preparation facilitates perceptual identification of letters. *Perception & Psychophysics*, 70(7), 1305-1313. https://doi.org/10.3758/PP.70.7.1305

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8), 1293-1313. https://doi.org/10.3758/bf03194544

Empirical Study on Auditory Perception in Autistic and Non-Autistic Hearers - Interplay between Prosody, Uncertainty, and Social Cognition

Charlotte Bellinghausen, Bernhard Schröder, Paula Dahmen: Institute of German Studies, University of Duisburg-Essen
Reinhold Rauh, Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg
Andreas Riedel: Department of Psychiatry and Psychotherapy, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg & Luzerner Psychiatrie, Ambulante Dienste, Löwengraben 20, 6004 Luzern
Peter Birkholz: Institute of Acoustics and Speech Communication, Technische Universität Dresden
Ludger Tebartz van Elst, Thomas Fangmeier: Department of Psychiatry and Psychotherapy, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg

Introduction: In our study, we tested the recognition of different degrees of intended uncertainty conveyed by prosodic cues as well as pitch discrimination in adult hearers with and without autism spectrum disorder (ASD). A systematic overview by Velikonja et al. (2019) showed large impairments in theory of mind (ToM) as well as emotion perception and processing in social cognitive domains for autistic hearers. According to Schwenck & Ciaramidaro (2014: 7), problems in recognizing specific emotions were often observed rather than a general emotion recognition impairment in ASD. Andres-Roqueta and Katsos (2020) found that children with ASD have problems with social pragmatics tasks involving ToM. It is not clear from the mentioned studies which role uncertainty recognition plays as social pragmatic competence when comparing hearers with ASD and neurotypical adults (NTC).

Hypothesis and Method: Since autistic individuals have difficulties in many areas of social cognition (for an overview of the role of social cognition in ASD see Schwenck & Ciaramidaro, 2014), it can be hypothesized that prosodic indicators of uncertainty have a weaker effect on uncertainty recognition in adult hearers with ASD compared to the NTC. We generated synthetic speech by using the articulatory speech synthesizer VocalTractLab (VocalTractLab, 2017): In that way we generated answers in a human-machine scenario with variations of *pause*, *intonation*, and *hesitation* for conveying different degrees of intended uncertainty. In the prosody task, 28 adults with ASD and 28 NTC rated each synthetic answer on a 5-point Likert scale regarding uncertainty, naturalness, and understandability. To control the intervening variable of differences in pitch perception between the groups, we carried out tests for the perception of discrete and gliding pitch changes.

Results: For uncertainty perception, our data show significant differences between the different degrees of intended uncertainty irrespective of the study groups. When comparing the groups, we found a tendency that listeners with ASD needed longer response times for uncertainty judgments at least in some conditions. For the combination *hesitation & strong intonation* the difference was significant. Furthermore, the different levels of uncertainty were in general rated more certain than in the NTC group, although differences did not survive correction for multiple testing. Furthermore, in the pitch discrimination and pitch change task no significant group differences were found.

Discussion: In summary, we found that descriptively the ASD group showed higher certainty ratings and slower response times in nearly all conditions, but only for the combination of hesitation & strong intonation group differences were significant. Possible explanations for the small to moderate effects could be that compensatory strategies are more easily acquired in prosodic uncertainty perception than in other areas of implicit social cognition tasks such as facial emotion recognition. Alternatively, autistic individuals may have more difficulties in uncertainty recognition in complex and natural settings. In future work, we would like to further investigate the interplay between prosody, social and non-social cognition, and auditory perception in autistic and non-autistic hearers.

*Keywords***:** autism spectrum disorder, prosody, auditory discrimination, social cognition

References

Andrés-Roqueta, C., Katsos, N. (2020). A distinction between linguistic- and social-pragmatics helps the precise characterisation of pragmatic challenges in children with Autism Spectrum Disorders and Developmental Language Disorder. *Journal of Speech, Language and Hearing Research*, 63(5), 1494-1508 https://doi.org/10.1044/2020_JSLHR-19-00263

Schwenck, C., Ciaramidaro, A., (2014). Soziale Kognition bei Autismus-Spektrum-Störungen und Störungen des Sozialverhaltens. *Kindheit und Entwicklung*, 23(1), 5-12.

Velikonja, T., Fett, A.-K, Velthorst, E. (2019). Patterns of nonsocial and Social Cognitive Functioning in Adults with Autism Spectrum Disorder. *JAMA Psychiatry*, 76(2),135-151. doi:10.1001/jamapsychiatry.2018.3645

Vocal Tract Lab (2017). www.vocaltractlab.de [last access: 22nd of july 2022]

Relating ecological accounts of cognition with the reinforcement learning

Olgierd Borowiecki, Faculty of Philosophy and Social Sciences, Nicolaus Copernicus University in Toruń, Poland

In this theoretical work, two frameworks describing acting in dynamic environments are related. The first framework is an ecological approach combined with evolutionary biology and embodied cognitive science. The second framework is reinforcement learning, a leading computational model in neuroscience of action and decision making.

In the first framework, a living organism and its surroundings are described as a coupled system filled with dynamic relations facilitating actions - affordances. An organism interacts with its environment, actively constructs a niche, which in turn guides an organism's actions. A niche therefore effectively reduces degrees of freedom that an organism-surroundings system has.

In the second framework, an environment is characterised in terms of states and possible transitions between the states. Such structure of an environment is called a Markov Decision Process environment (MDP). The reinforcement learning was initially developed for artificial agents, but subsequently it has been applied to animal and human studies. Specific tasks used for research on living organisms, (e.g. multi-step decision taks) have a structure which meets criteria for the MDP environment. An agent navigating such environment has to balance exploration and exploitation in order to maximise its utility function (collected rewards).

It is argued that relating the two frameworks allows for precise operationalisation of the environment and the niche - terms which are often used equivocally in the first framework. The environment is understood as a part of the organism-surroundings system which is explored by following affordances; the niche as a part which is exploited by following signifiers. The second framework is enriched with a naturalistic description of an agent's surroundings which in turn could be used for a better conceptualisation of a utility function. Progress in this direction is necessary for correlating brain activity with action and decision making performed by living organisms in naturalistic environments.

Keywords: Reinforcement Learning, Ecological Psychology, Evolutionary Biology, Affordances, Exploration-Exploitation Dilemma.

References

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. Synthese, 195(6), 2417-2444.

Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. Frontiers in human neuroscience, 8, 599.

Chemero, A. (2011). Radical embodied cognitive science. MIT press.

Clark, A., & Chalmers, D. (1998). The extended mind. analysis, 58(1), 7-19.

Co-Reyes, J. D., Sanjeev, S., Berseth, G., Gupta, A., & Levine, S. (2020). Ecological reinforcement learning. arXiv preprint arXiv:2006.12478.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature neuroscience, 8(12), 1704-1711.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. Neuron, 80(2), 312-325.

Di Paolo, E., & Thompson, E. (2014). The enactive approach. The Routledge handbook of embodied cognition, 68-78.

Flach, J. M., Stappers, P. J., & Voorhorst, F. A. (2017). Beyond affordances: Closing the generalization gap between design and cognitive science. Design Issues, 33(1), 76-89.

Frankenhuis, W. E., Panchanathan, K., & Barto, A. G. (2019). Enriching behavioral ecology with reinforcement learning methods. Behavioural Processes, 161, 94-100.

Gibson, J. J., & Carmichael, L. (1966). The senses considered as perceptual systems (Vol. 2, No. 1, pp. 44-73). Boston: Houghton Mifflin.

Gibson, J. J. (2014). The ecological approach to visual perception: classic edition. Psychology Press.

Lewontin, R. (2011). The genotype/phenotype distinction. Stanford Encyclopedia of Philosophy.

Norman, D. (2013). The design of everyday things: Revised and expanded edition. Basic books.

Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: feedback control as a process model of behavior and cognition. Trends in cognitive sciences, 20(6), 414-424.

Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. Ecological psychology, 26(4), 325-352.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

Varela, F. J., Thompson, E., & Rosch, E. (2017). The embodied mind, revised edition: Cognitive science and human experience. MIT press.

Disentangling population stereotypes from individual differences in ACEs

Francesca Capuano[1], Berry Claus[2], Daniel Bub[3], & Barbara Kaup[1]
[1]University of Tübingen; [2]Leibniz University Hannover; [3]University of Victoria

According to embodied-simulation theories of language processing, texts are comprehended by simulating the actions and referents they describe through the reactivation of sensorimotor traces. These theories are supported by *action-sentence compatibility effects* (ACEs): physical actions are performed faster when they match the characteristics of the actions described in a sentence. For example, when reading a sentence about a hand rotation movement, subjects rotate their hand faster in the direction of motion that matches the direction described (Zwaan & Taylor, 2006; but see Morey et al., 2022 for skepticism on the ACE).

A recent manuscript by Teskey et al. (2022) highlights how ACEs are not sufficient to support strong embodiment: most experiments so far cannot distinguish whether these effects stem from sensorimotor traces or simply from **population action stereotypes**, that is, widely shared ideas on how actions need to be performed in order to achieve a certain goal (e.g. most people expect to turn a volume knob clockwise to increase the volume). Most of the times, in fact, these action stereotypes match the actual corresponding motor experience. Teskey et al. (2022) exploit a case where - at least in English - direction of motion and population stereotype mismatch, in order to identify the influence of population stereotypes on ACEs: opening screw-top lids is stereotypically associated to a clockwise motion, but is actually performed counterclockwise. They find that language comprehension is sensitive to abstract representations such as action stereotypes. Teskey et al. attribute inconsistencies in the embodiment literature to foreseeing this aspect: Zwaan & Taylor (2006) found a classic ACE with sentences describing varied hand rotation actions, but Claus (2015) found a negative ACE as she presented only sentences that described the opening or closing of screw-top lids (corresponding exactly to the aforementioned mismatch between motor experience and action stereotype). A different explanation to the same inconsistencies is offered by Capuano et al. (2022), which highlight how individual differences could instead be responsible for divergent results. They provide evidence that ACEs might still stem from the individual reactivation of sensorimotor traces, but the reported inconsistencies might reflect differential preferences that individuals have with respect to which hand they typically employ on top of the lids vs. which hand they use on the containers. Differential preferences would give rise to opposing patterns of hand motion. The study was run in German.

As a contribution to identify the true source of ACEs, we tested the replicability of the *clockwise-to-open* stereotype found by Teskey et al. (2022) on 20 German subjects. In each trial, they saw one of two verbs (*öffnen* - Eng. *open* and *schließen* - Eng. *close*), together with the gifs of two knobs (one rotating clockwise and the other counterclockwise). They were instructed to click on the clockwise knob when they saw the verb *öffnen*, and on the counterclockwise knob when they saw the word *schließen* (or vice versa). The position of the knobs was counterbalanced. An analysis of RTs shows an interaction of verb and instruction (supporting the presence of the stereotype in German), which is not observed when employing different verbs (e.g. *niesen* - Eng. *sneeze* and *husten* - Eng. *cough*). Following Zwaan & Taylor (2006)'s evidence that visual stimuli also interact with the comprehension of sentences describing manual actions, this online replication extends the results to visual information and paves the way to more efficient further testing of the two concurrent proposals.

Keywords: Embodied cognition, motor stereotypes, language processing, online experiment.

References

Capuano, F., Claus, B., & Kaup, B. (2022). The experiential basis of compatibility effects in reading-by-rotating paradigms. *Psychological Research*, 1–13.

Claus, B. (2015). Verb gapping: An action-gap compatibility study. *Acta psychologica*, *156*, 104–113.

Morey, R. D., Kaschak, M. P., Díez-Álamo, A. M., Glenberg, A. M., Zwaan, R. A., Lakens, D., ... others (2022). A pre-registered, multi-lab non-replication of the action-sentence compatibility effect (ace). *Psychonomic bulletin & review*, *29*(2), 613–626.

Teskey, M. L., Bub, D. N., & Masson, M. E. J. (2022). Direction-of-motion stereotypes influence action execution and understanding of action description. (Manuscript submitted for publication)

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1.

Competitive memory operations during discourse comprehension

Susanne Dietrich*, Verena C. Seibold, & Bettina Rolke

Evolutionary Cognition, Department of Psychology, University of Tübingen

Discourse comprehension depends on an individual's ability to integrate new information into a given context. Understanding a discourse is difficult when presuppositions (PSPs), i.e., essential context information (Schwarz, 2016), are not given or are violated. PSPs are triggered by specific words signalling that a reference process to contextual information must be made. For example, the definite determiner "the" triggers the presupposition that there exists an item in the context to which a reference is needed (existence PSP). A discourse is not plausible when the reference process fails to find a referent in the context, i.e., when the referent is negated (violated existence PSP). The indefinite determiner "a" triggers the PSP that a context item is newly introduced and was previously non-existent (novelty PSP). This PSP is violated when an item in the context already exists (violated novelty PSP). Recent fMRI-data suggest that such PSP triggered reference process to contextual information require working memory demands, which increase when presuppositions are violated (Dietrich et al., 2019).

In this online-study, we investigated whether reference to PSP triggered information is interlinked with other memory processes during discourse understanding. To this end, we presented auditory sentences, and actively recruited naïve participants (N = 28) were asked to judge the coherence of a test sentence to a context which consist of four sentences and was presented before by using a four-point scale („How well does the test sentence match to the context?"). In sentence set (1), we manipulated working memory load by presenting the contextual referent for the definite determiner in the test sentence at far or close temporal distance within the context (Table 1, left column) (see O'Brien, 1987). To monitor PSP reference processes, we either fulfilled the existence PSP by mentioning the referent for the definite determiner or we violated the PSP by negating the potential referent (Coherence: intact vs. violated). In sentence set (2), a referent was not explicitly mentioned in the context, but inference of a referent was either plausible or implausible by contextual semantic relations (Table 1, right column). PSP processing in this sentence set was manipulated by presenting the indefinite determiner to introduce a new item to a context or by presenting the definite determiner, thereby violating the novelty PSP (Coherence: intact vs. violated). The two sentence sets of 64 trials each were presented in randomized order in a within-subject design. If PSP reference processes interfere with memory processes, we expected to see differences in PSP coherence judgements (difference between intact – violated conditions) depending on memory load and on plausibility of context.

Coherence judgements for each sentence set were analysed via separate linear mixed-effects models (LMEMs) using restricted maximum likelihood fitting. We fitted LMEMs containing (i) by-subjects random intercepts and slopes for Coherence and Distance or Coherence and Plausibility and their interaction and (ii) by-stimulus random intercepts. We report Type III *F*-tests with Satterthwaite-approximated degrees of freedom.

Figure 1 shows mean rating values depending on Coherence for sentence set (1, left) and sentence set (2, right). For sentence set (1), Coherence, $F_{Sat}(1, 26.99) = 118.36$, $p < .001$, as well as Distance, $F_{Sat}(1, 34.54) = 5.90$, $p = .020$, affected judgements. Importantly, the interaction of Coherence and Distance, $F_{Sat}(1, 28.96) = 18.18$, $p < .001$, shows that the Coherence effect was less pronounced for far distance items than for close ones. The Coherence effect was absent in sentence set (2), $F_{Sat}(1, 80.21) = 0.63$, $p = .430$, and there was no interaction between Coherence and Plausibility, $F_{Sat}(1, 27.00) = 0.46$, $p = .604$, but there was an effect of Plausibility, $F_{Sat}(1, 34.26) = 105.05$, $p < .001$.

Our results revealed a strong influence of working memory demands on PSP reference processes triggered by the definite determiner. If the context referent was far more distant and

most probably less clearly represented in memory, PSP violations were judged to be less severe than those for closer referents. Moreover, PSP violations seem to play no role when the semantic context provides a basis for a plausible presence of a referent. Our results suggest that discourse understanding involves short-lasting reference processes, which might fade in memory with increasing distance or can be overlaid by semantic contextual content.

Keywords: **presupposition, context-dependent speech, reference, coherence**

Table 1. Examples for the stimulus materials. In sentence set (1) (Remember context, left column) the reference information was either presented (intact) or not (violated) and it was presented at close or far distance. In sentence set (2) (Infer from context) the context was either plausible or implausible for a potential referent and the referent was mentioned by the indefinite determiner (intact) or by the definite determiner (violated).

(1) Remember context	**(2) Infer from context**
Intact - Far condition	Intact - Plausible condition
Context sentences: (1) Tina is going to France. (2) Tina has **a suitcase**. (3) The travel is short. (4) Tina has no duffel bag.	**Context sentences:** (1) Simon rearranges the living room. (2) Simon has a sofa. (3) He puts his things in different places. (4) Simon has no TV.
Test sentence: When Tina packs **the suitcase**, she notices a dark cloud.	**Test sentence:** When Simon moves **a shelf**, he suffers a painful strain.
Violated - Close condition	Violated - Implausible condition
Context sentences: (1) Tina is going to France. (2) Tina has a duffel bag. (3) The travel is short. (4) Tina has **no suitcase**.	**Context sentences:** (1) Simon goes hiking in the mountains. (2) Simon has a tent. (3) He looks forward to the trip for a long time. (4) Simon has no rope.
Test sentence: When Tina packs **the suitcase**, she notices a dark cloud.	**Test sentence:** When Simon moves **the shelf**, he suffers a painful strain.

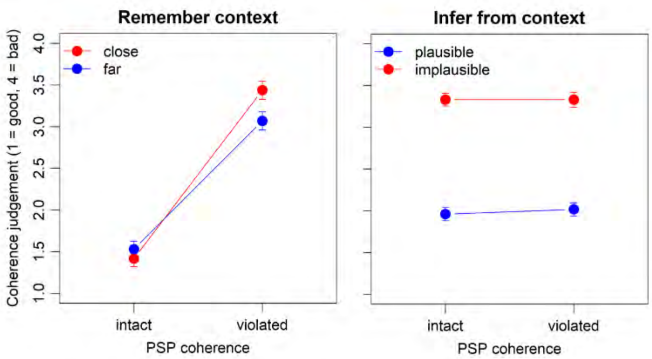


Figure 1. Mean values of coherence judgement as a function of PSP coherence. In sentence set (1) (Remember context, left panel) the reference information was presented at close or far distance while in sentence set (2) (Infer from context, right panel) the context was either plausible or implausible for a potential referent. Error bars represent +/- 1 standard error of the mean, being corrected for within-subjects design according to a method proposed by Morey (2008).

References

Dietrich, S., Hertrich, I., Seibold, V. C., & Rolke, B. (2019). Discourse management during speech perception: A functional magnetic resonance imaging (fMRI) study. *NeuroImage*, 202:116047. doi: 10.1016/j.neuroimage.2019.116047

Morey, R. D. 2008. Confidence Intervals from Normalized Data: A correction to Cousineau (2005). Tutorial in Quantitative Methods for Psychology 4: 61-64.

O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 278 – 290.

Schwarz, F., 2016. Experimental work in presupposition and presupposition projection. *Annual Review of Linguistics*, 2, 273–292. doi: 10.1146/annurev-linguistics-011415-040809.

Relating Enactive Sense-Making to Psychological Meaning-Making

Roy Dings (Ruhr University Bochum) & Caroline Stankozi (Ruhr University Bochum)

How is it that ordinary people typically experience the world as meaningful? Conversely, why do people suffering from depression experience the world as meaningless? Psychologists often investigate meaninglessness in depression in terms of 'meaning-making' processes (Park 2010). They locate meaning on the personal level, e.g. as "a mental representation of possible relationships between things, events and relationships. Thus, meaning *connects* things" (Baumeister 1991, p.15; original emphasis). In this philosophical paper we aim to show that the enactivist concept of 'sense-making' is able to supplement both research and therapy by adding a more narrow-grained (subpersonal) description of what happens in processes of meaning-making. This is particularly relevant for therapeutic contexts. A therapist treating depression has to be sensitive to personal-level problems (e.g. a lack of experienced meaning) but may rely on knowledge of subpersonal-level processes to address these problems. Therapy requires linking those levels – just as most research on meaning (even in AI) does.

According to De Haan (2020), an enactive approach to psychiatry is able to "integrate the physiological, experiential, sociocultural, and existential dimensions of psychiatric disorders" (ibid., p. 263). This enactive account "understands the mind as sense-making: as a type of interaction that emerges from organisms that depend on their environments" (ibid., p. 234). As soon as an organism senses an object, appraises it through bodily arousal, and responds to it in an adaptively regulated manner, sense is made within the coupled organism-environment system (Di Paolo, Buhrmann & Barandiaran 2017, p. 124). The normative evaluation (for adaptivity) happens in a self-regulatory fashion: Competing sensorimotor coordination patterns get excited and inhibited by each other (and by further, e.g. preceding, patterns). The pattern that ends up with the most support then "dominates" the others and – bodily and environmental conditions allowing – is enacted (ibid., pp.189, 200). How does this subpersonal-level approach relate to psychological work on personal-level meaning-making?

To clarify this matter we take our lead from De Haan's (2020) suggestion to understand (disruptions of) enactive sense-making in terms of (alterations in) an agent's field of affordances. She writes that "sense-making discloses a world, which is to say: not a neutral collection of objects, but rather a world that is significant to me, offering me various possibilities for action [i.e. affordances]" (ibid., p. 218). For a depressed person, however, the disordered sense-making leads to a field of affordances that is devoid of meaning (ibid., p. 219, Fig. 7.1). Here we think it is useful to distinguish between the *relevance* (for mere sense-making) and the *meaningfulness* of an affordance (cf. Dings 2021). An affordance is *relevant* when it matters to an agent in the immediate context of the action. For instance, a chair is relevant to us if we want to sit down. An affordance is *meaningful* if it is relevant to the agent across various (temporally extended) contexts. Whether the chair is broken or not, it might still be the one your grandfather crafted himself – and thereby mean something to you. For severely depressed persons, things typically loose at least some of their meaningfulness. The chair may no longer mean anything to them, aside being relevant for sitting down.

De Haan's (2020) enactive account of sense-making seems to accommodate only the *relevance* of affordances. This makes it insufficient to capture depressive experience, as depressed people may certainly experience relevance. While her notion of 'existential sense-making', which concerns making sense of one's sense-making, goes in the right direction, it does not seem to account for navigating cross-temporal contexts. An account of such *meaningfulness* is precisely what the psychological work on meaning-making provides. Navigating temporally extended contexts, and establishing *relations* of relevance, is what is central to meaning-making processes like self-narration (Singer 2004). Self-narration may contribute to the modulation of bodily sensitivity to particular affordances (Dings 2019, p. 20), suggesting a dynamic interplay of sense-making and meaning-making.

*Keywords***:** Sense-Making, Meaning, Depression, Affordances, Enactivism

References

Baumeister, R. F. (1991). *Meanings of Life*. Guilford Press.

De Haan, S. (2020). *Enactive Psychiatry*. Cambridge University Press.

Di Paolo, E., Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor Life*. Oxford University Press.

Dings, R. (2019). The dynamic and recursive interplay of embodiment and narrative identity. *Philosophical Psychology, 32*(2), 186–210.

Dings, R. (2021). Meaningful affordances. *Synthese, 199*, 1855–1875.

Park, C. L. (2010). Making sense of the meaning literature: an integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological bulletin*, 136(2), 257.

Singer, J. A. (2004). Narrative identity and meaning making across the adult lifespan: An introduction. Journal of personality, 72(3), 437–460.

On the processing of nominal compounds: the role of context and familiarity

Francisco M. I. Gómez[1], John C. B. Gamboa[1], Juhani Järvikivi[2], Shanley E. M. Allen[1]
[1]Technische Universität Kaiserslautern, [2]University of Alberta

The Uniform Information Density (UID) hypothesis (Jaeger, 2010) views communication between speakers as transmission of information through a noisy channel. It asserts that, in order to achieve efficient and reliable communication (under the perspective of Information Theory, cf. Shannon, 1948), speakers strive to transmit information uniformly at the capacity of this channel, avoiding peaks of information. If the capacity is exceeded, it predicts that the receiver is to experience processing difficulty. The information carried by the words depends on their probabilities: the more probable a word is, the less information it carries.

Nominal compounds (NCs; e.g., *morphine reward effect*) are structures able to compress into few words complicated concepts (compare, *the effect of morphine on reward mechanisms*), and could arguably constitute peaks of information. As such, the UID hypothesis predicts they should not be common. Surprisingly, however, they are ubiquitous in scientific texts, being sometimes hard to understand in isolation (e.g., *start arm barrier*) but also sometimes quite easy (e.g., *stress response gene*). This frequency may be explained in a number of ways. Perhaps the context "sets up" the NC meaning before it is used, making NCs more expected, raising their probability, thus reducing their information content; or perhaps the NC is already quite familiar to the reader. In this work, we investigate the role of these two variables in NC processing.

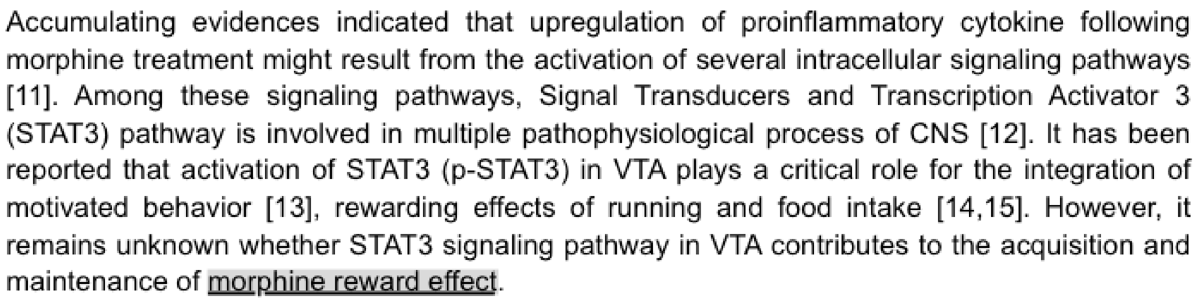
In the experiment below, items were text passages containing NCs extracted from real scientific papers in the areas of Economics, Biology and Linguistics (see Figure 1). NCs were either *Low* or *High* familiarity, as determined using a formula that took into account data from Wikipedia and from the Google Ngram Viewer (Michel et al., 2011). In addition, as a measure of contextual set up, we counted the number of words in the context preceding the NC that were similar to the words composing the NC (where the degree of similarity was calculated using the cosine similarity of BERT word vectors; Devlin et al., 2018) and categorized NCs into having either *Few* or *Many* similar words in its context.

Participants (N=56) read 24 text passages containing NCs, and 24 NCs in isolation, answering in a 1-10 scale (1: very easy; 10: very ard) how much difficulty they perceived to understand the NC, and producing a paraphrase of their meaning. We expected the difficulty of familiar items to be lower than that of unfamiliar items for both items in isolation and items in context. In addition, strong set up text passages (with *Many* similar words) should evince less difficulty than weak set up passages (with *Few* similar words); but this effect should not be present for NCs presented in isolation.

For our preliminary analysis, we analysed separately the results for NCs presented in context and in isolation (see Figure 2), fitting a separate linear mixed-effects model for each dataset. It can be seen that the data are quite similar for both conditions. We found no significant effect of set up strength. However, we did find an effect of familiarity, indicating that familiar items were judged easier both for NCs presented in context ($t = 7.546$, $p < .001$) and in isolation ($t = 7.095$, $p < .001$).

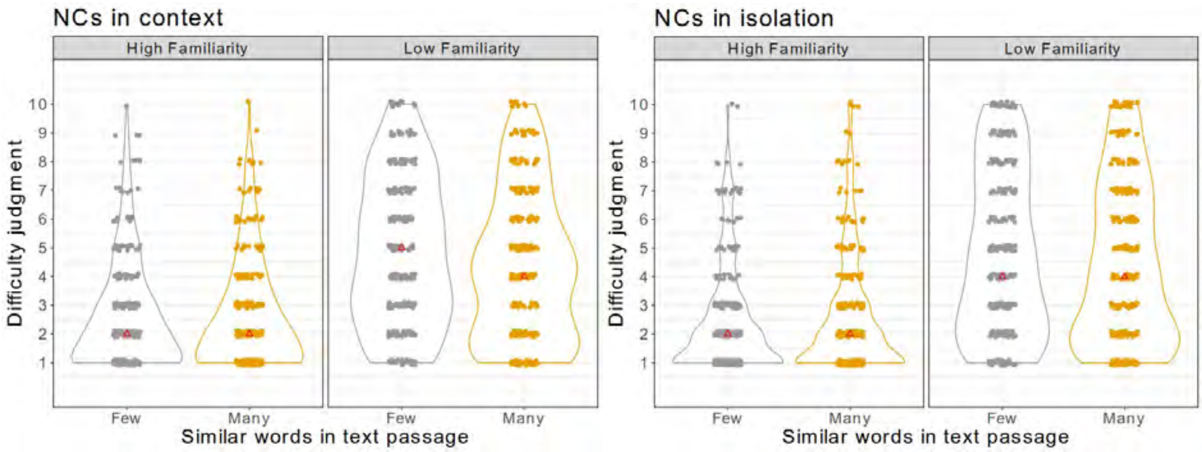
Despite having found no effect of contextual set up, it is not possible to conclude that NCs are not set up in their context, or what the relation between the UID hypothesis and NC use is. Our results may be explained in a number of ways. For example, it is not clear how well the BERT word vectors used to determine word similarity reflect the judgments made by humans. In addition, it is possible that our measure of contextual set up, which did not take into account potentially relevant information such as sentence structure, was not sensitive enough for this task. In future studies, we will investigate these explanations.

Keywords: nominal compounds, context, familiarity, sentence processing, information density



Accumulating evidences indicated that upregulation of proinflammatory cytokine following morphine treatment might result from the activation of several intracellular signaling pathways [11]. Among these signaling pathways, Signal Transducers and Transcription Activator 3 (STAT3) pathway is involved in multiple pathophysiological process of CNS [12]. It has been reported that activation of STAT3 (p-STAT3) in VTA plays a critical role for the integration of motivated behavior [13], rewarding effects of running and food intake [14,15]. However, it remains unknown whether STAT3 signaling pathway in VTA contributes to the acquisition and maintenance of morphine reward effect.

Figure 1. Example text passage as it was presented in the questionnaire.



(a) (b)

Figure 2. Violin plots of the participant judgments for NCs presented (a) in context and (b) in isolation. The red triangle is indicates the median of the responses.

References

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*(1), 23-62.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Google Books Team, ... & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331*(6014), 176-182.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*(3), 379-423.

The Effect of Feedback and Motivation on Time Estimation Performance

Stine Hollah*, Gerke Feindt*, & Sebastian Wallot

Institute for Sustainability Education and Psychology, Leuphana University, Germany

* These authors contributed equally to the present study and manuscript.

Introduction

The current study examined the effects of motivation and visual accuracy feedback on participants' performance in a time estimation task. Earlier studies investigating feedback time estimation – counterintuitively – found no effect of such feedback on standard deviation of time intervals (Kuznetsov & Wallot, 2011). Moreover, to date nothing is known about the influence of motivational factors on time estimation. This is, however, of importance because participants often experience time estimation tasks as relatively "boring". To re-examine the effect of feedback on dispersion of time intervals, participants produced time intervals either with or without accuracy feedback in the lab. Next to the standard deviation of intervals, we also examine the absolute deviation from the target interval as a measure of accuracy that includes participants' inter-individual differences in preferred time estimates (McAuley et al., 2006). To examine motivational effects, one group of participants received a monetary incentive if their time intervals were close to the target interval of one second (Mir et al., 2011).

Methods

Participants were students at the Leuphana University Lüneburg, $N = 21$ (57.1% females, 19-51 years, mean = 22.9 yrs., $SD = 6.7$ yrs.). They were compensated with course credits or cash. Informed consent was obtained from all participants. Within the OpenSesame (Mathôt, 2010) based timing experiment, participants were instructed to press the space-key of the experimental laptop (Dell, Intel® Core™ i5-9400H CPU, 2.50GHz, 8.00 GB) repeatedly in 1-s intervals directly after listening to 10 beats of the 1-s interval. Participants performed the time estimation task without and with visual feedback (within participant factor, order counterbalanced), showing them after every tap by how many milliseconds they missed the target interval. Half of the participants received a monetary incentive of 1 Eurocent for each time interval that deviated less than 100ms from the target interval. As participants had to produce 1100 intervals per session, this could lead to a maximum payoff of 11 EUR. Finally, participants were asked 21 questions about motivation, difficulty and demographics.

Results

Mixed ANOVAs did not show effects of feedback or motivation on SD of time intervals (all ps > .37) - (Figure 1a), even when definite outliers were removed (all ps > .07) - (Figure 1b). A significant main effect of receiving feedback on the absolute deviations from the 1-s interval ($F(1, 19) = 25.56$, $p = .00007$) – (Figure 2), but no effect of motivation on timing performance was apparent ($F(1, 19) = 0.11$, $p = .742$).

Discussion

The present study shows that accuracy feedback on time estimation leads to an increase in accuracy of time intervals when the absolute deviations from the target interval are considered. However, we did not find effects of motivation by monetary incentive on time estimation although motivation within the incentive condition was higher when compared to the no incentive condition ($F(1, 16) = 24.806$, $p < .001$). Past research has shown that monetary incentives improve reaction times (Mir et al., 2011), suggesting participants might lack the motor-capacity to increase their performance, or that the effect is substantially smaller compared to that of accuracy feedback. Further questionnaire variables are not yet evaluated.

Keywords: time estimation, accuracy feedback, performance, motivation, monetary incentive

References

Kuznetsov, N., & Wallot, S. (2011). Effects of accuracy feedback on fractal characteristics of time estimation. *Frontiers in integrative neuroscience*, *5*, 62. Doi: 10.3389/fnint.2011.00062

Mathôt, S., (2010). *Opensesame* (3.3.10) [Software]. Mathôt, S. https://osdoc.cogsci.nl/

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348-367. Doi: 10.1037/0096-3445.135.3.348

Mir, P., Trender-Gerhard, I., Edwards, M. J., Schneider, S. A., Bhatia, K. P., & Jahanshahi, M. (2011). Motivation and movement: the effect of monetary incentive on performance speed. *Experimental brain research*, *209*(4), 551-559. Doi: 10.1007/s00221-011-2583-5

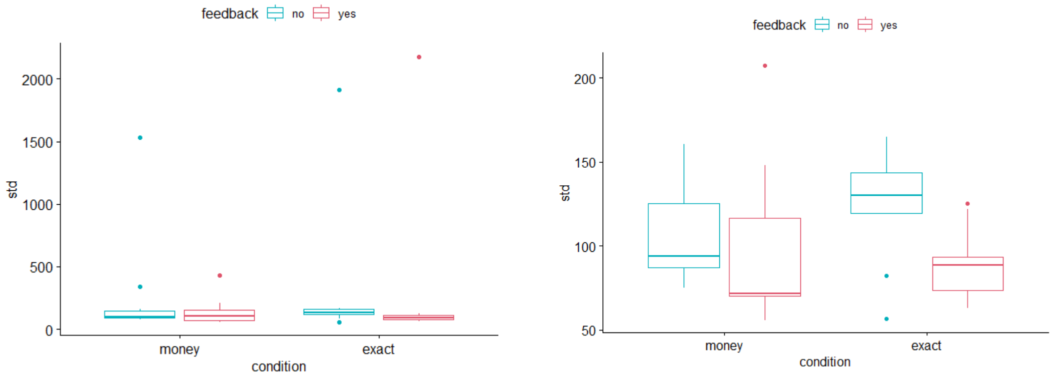


Figure 1a (left). Standard deviations of time estimates as a function of the presence of accuracy feedback (with or without feedback) combined with monetary incentive (only exact feedback or feedback combined with monetary incentive). Neither feedback, nor monetary incentive led to an increase in accuracy of time estimates. *Figure 1b (right).* Repetition of analysis while excluding definite outliers according to the criterium of ± 1.5 IQR.

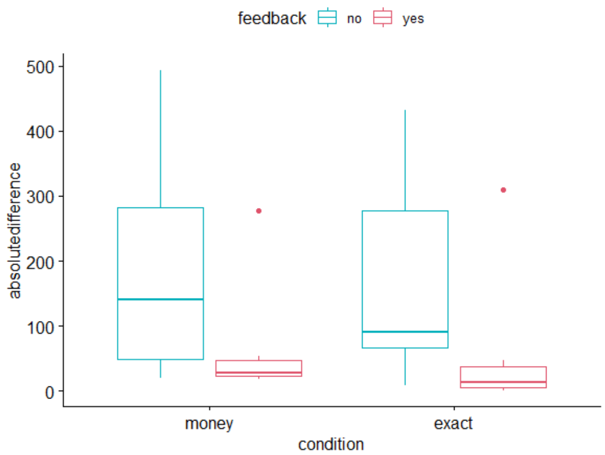


Figure 2. Absolute deviations of time estimates from the target interval (1s) as a function of the presence of accuracy feedback (with or without feedback) combined with monetary incentive (only exact feedback or feedback combined with monetary incentive). Feedback, but not monetary incentive led to an increase in accuracy of time estimates.

Speaking Together versus Speaking Alone: A Mega-Analysis of Six Experiments on Joint Language Production

Holtz, Nora[1], Hauber, Roger[1], Abdel Rahman, Rasha[1,2], Kuhlen, Anna K.[2]

[1]Berlin School of Mind and Brain, Berlin, Germany
[2]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

We conducted a mega-analysis of six experiments conducted in our lab (Hauber et al., unpublished; Kuhlen & Abdel Rahman, 2017, 2021). The studies investigate cumulative semantic interference due to categorical relations, an increase in naming latencies participants typically show when continuously naming a series of semantically related pictures (e.g., Howard et al., 2006). Importantly, to probe theories that view language production as a joint action where a task partner's speaking can influence individual's own speaking, the authors implemented this paradigm in a joint task setting in which participants take turns naming pictures together with a task partner. The original experiments had come to conflicting results regarding the existence of partner-elicited semantic interference which is operationalized as an increase in naming latencies induced by related pictures the task partner names.

We applied linear mixed models to the pooled data set to reconcile findings across the experiments and to explore the data for potentially moderating effects that could explain the divergent findings. Across models, two effects emerged: (1) All models revealed a significant main effect of Naming Order, reflecting an increase in naming latencies with each additional picture participants named within a given semantic category. This effect replicates the typical cumulative semantic interference effect. Crucially, (2) the models also revealed a small but robust interaction effect of Naming Order and Naming Condition, indicating that cumulative semantic interference was stronger in the condition where in addition to the participant themselves, the task partner named semantically related pictures. This finding is in line with the assumption that speaking together with a partner can have lasting influences on the own language production system and supports findings from previous studies on picture naming in a joint task setting (e.g., Baus et al., 2014; Hoedemaker et al., 2017).

Explorative analyses revealed that the amount of this partner-elicited interference is likely to be affected by the interplay of multiple factors, which together determine how salient participants perceive their task partner to be and to what degree they perceive the setting of the common task to be social.

The finding of a small but robust partner-elicited interference effect in our mega-analysis gives evidence for the claim that speaking together is different from speaking alone, encouraging the design of further studies that capture the social nature of language production. Methodologically, we demonstrate how synthesizing studies in a mega-analysis based on original raw data allows researchers to perform sophisticated statistical analysis of even a small number of pooled experiments that would not be possible with conventional meta-analysis based on average data.

Keywords: mega-analysis, language production, joint action, semantic interference, linear mixed models

Figure 1. Mean naming latencies (in milliseconds) for each ordinal position in the Joint Naming condition and the Single Naming condition. Naming latencies are averaged across experiments, participants, and stimuli.

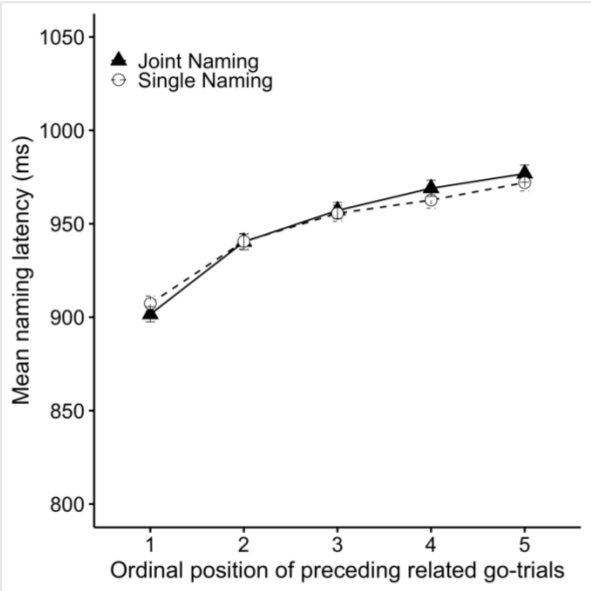


Table 1. Output of the selected Linear Mixed Model for the fixed effects structure.

Formula	*log(RT) ~ experiment*naming condition*naming order + (naming condition + naming order | subject) + (experimt*naming order | stimulus)*					
	Log-Transformed Target Naming Latencies					
	Sums of squares	Mean squares	Numerator degrees of freedom	Denominator degrees of freedom	F-value	p-value
Experiment	2.4081	0.4816	5	174.4	9.767	<0.001 ***
Naming Condition	0.1813	0.1813	1	319.1	3.677	0.05618 .
Naming Order	13.4768	13.4768	1	301.6	273.291	<0.001 ***
Experiment: Naming Condition	0.4228	0.0846	5	194.6	1.715	0.1329
Experiment: Naming Order	1.4760	0.2952	5	286.7	5.986	<0.001 ***
Naming Condition: Naming Order	0.3729	0.3729	1	32152.7	7.562	0.006 **
Experiment:Naming Condition: Naming Order	0.3679	0.0736	5	30875.0	1.492	0.1886

*. p<0.1 * p<0.05 ** p<0.01 *** p<0.001*

F-test statistics and corresponding p-values are based on the Satterthwaite approximation of the denominator degrees of freedom (Kuznetsova et al., 2017).

References

Baus, C., Sebanz, N., Fuente, V. de la, Branzi, F. M., Martin, C., & Costa, A. (2014). On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition*, *133*(2), 395–407. https://doi.org/10.1016/j.cognition.2014.07.006

Hauber, R., Kuhlen, A. K., & Abdel Rahman, R. (unpublished).

Hoedemaker, R. S., Ernst, J., Meyer, A. S., & Belke, E. (2017). Language production in a shared task: Cumulative Semantic Interference from self- and other-produced context words. *Acta Psychologica*, *172*, 55–63. https://doi.org/10.1016/j.actpsy.2016.11.007

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Kuhlen, A. K., & Abdel Rahman, R. (2017). Having a task partner affects lexical retrieval: Spoken word production in shared task settings. *Cognition*, *166*, 94–106. https://doi.org/10.1016/j.cognition.2017.05.024

Kuhlen, A. K., & Abdel Rahman, R. (2021). Joint language production: An electrophysiological investigation of simulated lexical access on behalf of a task partner. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(8), 1317–1337. https://doi.org/10.1037/xlm0001025

“Flashing upon that inward eye”
A neurocognitive poetics approach to mental imagery

Katharina Gloria Hugentobler & Jana Lüdtke
Freie Universität Berlin

Why do we “see” the things we see when we read a poem? Reading and understanding poetic texts is often described as an interactive process influenced by the words and phrases building the poems and all associations and images induced by them in the reader’s mind (Rosenblatt, 1978). We aim to shed light on the effect of words constituting a poem and their associations on the affective and cognitive processes while reading it. Our research is based on the Neurocognitive Poetics Model (Jacobs, 2015) which is conceptualised as dual-process model informed by Gestalt theory. As such it distinguishes between back- and foreground elements and postulates different influences on fast and more automatic compared to slow and more controlled comprehension processes. Moreover, the understanding of a poem is conceived as successful closing of a good Gestalt, *sensu* Iser (1972). Following this line of thought, we investigated the effect that semantic cohesion, i.e., the internal connection of a list of words operationalized via semantic similarity measures, has on understanding and appreciation of poetic texts and contrasted that to a word frequency manipulation. To do this, word lists varying regarding their internal semantic cohesion and mean word frequency were presented as modern micropoems (cf. Figure 1) to native English speakers. The (ease of) extraction of associated concepts as well as the affective and aesthetic responses are implicitly measured through asking for one-word titles and measuring reaction times, and explicitly measured with rating scales, e.g. for liking, induced mood, and ease of comprehension. The results presented here indicate that titles for micropoems composed of semantically related words were found more readily and seem to activate more unifying concepts indicated by less variations in the titles produced (cf. Figure 2). Moreover, semantically coherent micropoems were judged as more emotional and induced more appreciation. Comparable effects were not observed for word frequency manipulation. We interpret these results as evidence for the assumed relationship between generation of spontaneous associations, formation of mental imagery, and understanding and appreciation of poetic texts. Alternative explanations, for example based on ease of processing, seems to be less plausible as no effect of word frequency was observed. To further explore the role of semantic associations on mental imagery we currently replicate and extend our approach using German micropoems. In doing so we manipulated imageability of the poem words in addition to semantic cohesion and created a 2x2 (imageability x cohesion) poem matrix designed to investigate the individual and combined effects of both variations on our direct and indirect measures.

Keywords: neurocognitive poetics, associations, literary reading, text comprehension, computational linguistics

result	transplant	bed
banana	beat	dream
fierce	love	awake
active	soul	night
tape	ache	rest
cigarette	throb	peace
equality	attack	sound
poor	pump	tired
charm	red	death

Figure 1. Three examples of items presented as micropoems with increasing (from left to right) semantic cohesion

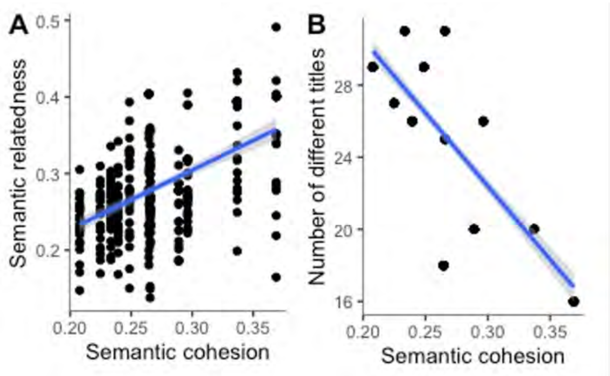


Figure 2. Implicit measures: Relationship between semantic cohesion and semantic relatedness between titles and poems (A) and number of different titles (B).

References

Iser, W. (1972). The Reading Process: A Phenomenological Approach. *New Literary History*, 3(2), 279–299.

Jacobs, A. M. (2015). Neurocognitive poetics: Methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, 9, 1–22.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: the transactional theory of the literary work*. Carbondale: Southern Illinois University Press.

Does deliberate prospection help students set better goals?

Jähnichen, Sarah[1]; Weber, Felix[1]; Prentice, Mike[2]; Lieder, Falk[2]
[1]Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany
[2]Rationality Enhancement Group, MPI for Intelligent Systems, Tübingen, Germany

Human cognition is fundamentally goal-directed (Carver & Scheier, 2001), and there are still many open questions about the cognitive mechanisms of goal-setting and how they affect the quality of people's goals (Kasser & Ryan, 1996). Here, we study in an exploratory way how goals set through deliberate reflection about the future (prospection) differ from the goals people set without guided prospection. We conducted an online experiment where students derived goals from imagining what they would like a close friend or relative to say about their accomplishments in the future. We aimed to answer two questions: 1) How does this goal-setting strategy affect the goal's characteristics (e.g., importance and specificity), and which broad life aspirations the goal aims at (e.g., wealth, safety, or happiness)? 2) How do those effects depend on the time horizon of the imagined accomplishments? The second question is especially important because the impact of the time horizon of prospection has not been studied before. In the first phase of the experiment, all participants engaged in unstructured, baseline goal-setting. In the second phase of the experiment, all participants were guided to engage in deliberate prospection. Concretely, they imagined what they would like a close friend or relative to say about their accomplishments at one of three randomly assigned time points: the end of the current semester ($N = 26$), the end of their university studies ($N = 26$), or the end of their life ($N = 24$). Finally, they were asked to derive a goal from the resulting insights. Participants rated both goals on shortened and translated versions of the Goal Characteristics Questionnaire (Iwama et al., 2021) and the Aspiration Index (Kasser & Ryan, 1996; Grouzet, et al., 2005) immediately after setting each goal.

Regarding our first question, we found that the goals students set through deliberate prospection were superior to the goals they set intuitively on several metrics (see Table 1) but less specific and less measurable. Participants rated post-prospection goals as being more useful to pursue ($p<.001$), better aligned with their identity ($p<.01$) and their values ($p<.05$), more strongly tied to aspirations to have exciting life experiences ($p<.05$), and less strongly tied to wealth ($p<.05$) and safety ($p<.01$) aspirations. Together, these findings are consistent with the theory that people's attention has a positivity offset highlighting possible upsides when they think about psychologically-distant possibilities and a negativity bias highlighting possible dangers of more proximal goal pursuit (Caciappo & Bernston, 1999). Moreover, deliberate prospection shifted students' priorities towards maintaining what is good ($p<.05$) and pursuing worthwhile activities rather than focusing on the outcome ($p<.01$).

Regarding our second question about the time horizon of prospection, we found that reflecting on what they would like to have accomplished by the end of their life made students care less about financial success and more about contributing to their community than reflecting on shorter timescales (both $p<.05$). This finding indicates an increased focus on intrinsic values and a decreased focus on extrinsic values when reflecting on a longer time horizon. Our results also suggest that reflecting on what they would like to have accomplished by the end of their studies directs students more toward conformity-oriented goals than prospecting on other time periods ($p<.05$). This might be because conforming to academic norms and expectations is instrumental for getting a good job after graduation.

Overall, our findings are consistent with the hypothesis that deriving goals through deliberate prospection helps people set goals that they believe will be rewarding to pursue, and that research has shown are conducive to well-being. In some cases, these benefits to goal-setting are particularly pronounced when reflecting on the end of one's life. We will test several concrete hypotheses suggested by these results in a large, preregistered follow-up experiment. The aim is to characterize the benefits of deliberate prospection and to determine which time horizon is most beneficial for goal attainment and well-being.

*Keywords***:** goals; goal-setting; prospection; education; goal characteristics

Table 1. Effects of goal-setting (gs) and time horizon (t) on goal characteristics and aspirations. Statistically significant findings are highlighted in gray.

Measure	Unstructured	Deliberate prospection	p_{gs}	End of semester	End of studies	End of life	p_t
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	Mean (SD)	
Aspirations							
Hedonism	3.70 (1.16)	4.04 (1.04)	.019	3.89 (0.96)	4.12 (1.00)	4.13 (1.20)	.629
Money	4.29 (0.77)	4.03 (0.91)	.013	4.20 (0.75)	4.24 (0.72)	3.63 (1.14)	.047
Safety	4.07 (0.99)	3.66 (1.27)	.002	3.50 (1.34)	3.77 (1.31)	3.71 (1.17)	.584
Community	3.65 (1.08)	3.70 (1.09)	.689	3.43 (1.14)	3.58 (0.99)	4.13 (1.04)	.035
Conformity	1.89 (0.98)	1.95 (1.10)	.561	1.70 (0.89)	2.20 (1.24)	1.96 (1.13)	.025
Characteristics							
Utility of Goal Pursuit	3.85 (0.97)	4.18 (0.76)	.001	4.04 (0.78)	4.27 (0.78)	4.21 (0.73)	.644
Self-Congruence	3.14 (1.23)	3.50 (1.20)	.008	3.27 (1.08)	3.54 (1.31)	3.71 (1.20)	.592
Value-Congruence	3.65 (1.06)	3.94 (1.04)	.019	3.74 (1.08)	3.97 (1.04)	4.13 (1.00)	.502
Content Specificity	3.61 (1.30)	3.19 (1.28)	.017	3.43 (1.18)	2.93 (1.33)	3.21 (1.32)	.327
Time Specificity	3.36 (1.33)	2.95 (1.30)	.019	3.12 (1.18)	2.85 (1.32)	2.88 (1.43)	.506
Measurability	3.69 (1.08)	3.25 (1.24)	.002	3.31 (1.23)	3.00 (1.20)	3.46 (1.29)	.940
Maintenance-Attainment Goal	2.73 (1.30)	3.08 (1.35)	.036	2.97 (1.32)	3.39 (1.27)	2.88 (1.46)	.469
Process-Outcome Focus	3.73 (1.12)	4.12 (1.00)	.002	4.08 (1.02)	3.93 (1.10)	4.38 (0.83)	.055

References

Cacioppo, J. T., & Berntson, G. G. (1999). The Affect System: Architecture and Operating Characteristics. Current *Directions in Psychological Science, 8*(5), 133–137.

Carver, C. S., & Scheier, M. F. (2001). On the self-regulation of behavior. *Cambridge University Press*.

Grouzet, F. M., Kasser, T., Ahuvia, A., Dols, J. M., Kim, Y., Lau, S., . . . Sheldon, K. M. (2005). The structure of goal contents across 15 cultures. *Journal of personality and social psychology, 89*(5), 800.

Iwama, G. Y., Weber, F., Prentice, M., & Lieder, F. (2021). Development and Validation of a Goal Characteristics Questionnaire. *OSF Preprints*.

Kasser, T., & Ryan, R. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, *(22)*(3).

Role of Rehearsal in the Effects of Irrelevant Speech and Word Length

Abdullah Jelelati[1], Larissa Leist[1], Thomas Lachmann[1,2], Maria Klatte[1]

[1]Technische Universität Kaiserslautern, [2]Universidad Nebrija

Place-Name Categorisations Affect Wayfinding Behaviour

Lilian LeVinh,
Hanspeter A. Mallot
Eberhard Karls Universität Tübingen, Tübingen, Germany

Regionalisation is the grouping of places into regions. This can affect spatial reasoning in path planning as inter-regional distances tend to be overestimated (Burris & Branscombe (2005); Friedman & Montello (2006); Wiener & Mallot (2003)). Thus, when choosing the shortest route to a goal where two equidistant routes exist, humans prefer the route with less inter-regional border crossings (Region dependent route preference). While many cities have clusters of places or streets all bearing thematically grouped names, it is not clear whether the semantic grouping of place names can also evoke a perceived region. While it has been shown, that perceived regionalisation can stem from the semantic and contextual similarity of landmark objects (Schick et al. (2019), Wiener & Mallot (2003)), the same was not shown for contextual similarity of place names. Here we tested whether subjects show region dependent route preference according to semantically grouped place names. Further we tested the flexibility of such regionalisation by using place names with ambiguous relations, which fit into either of two equally logical categories. By presenting the place names in different contexts at the start of the experiment, subjects were biased to categorise place names differently. If route preference changed according to the previously discovered categorisation, it would support the hypothesis that regions used in wayfinding depend on the perceived semantic relations. We used a hexagonal ring maze with 12 places presented in virtual reality, as can be seen in Figure 1. Each place had three banners bearing the place name, otherwise all places were visually indistinguishable. Prior to the wayfinding task, subjects were asked to categorise the landmark names into three groups of four. By changing the context in which the place names were presented, subjects were primed to discover one or the other categorisation. Subject then explored the environment. In the test phase subjects were asked to take the shortest route between three places. As can be seen in Figure 2, the test tasks could be solved by taking either of two equidistant routes. The routes differed however in the number of inter-regional border crossings; depending on the previous categorisation one route will lead to more border crossings between regions than the other.
We analysed the categories formed by the subjects, five subjects were excluded as they formed at least one illogical or inconsistent category. In addition to two dropouts, we could thus use the data of 19 out of 26 subjects. As the subjects' goal was to take the shortest route possible, any unnecessary u-turns would result in a failure. Therefore the turn at the first decision point predetermines the route taken and was thus used to assess, whether route planning conformed with the previously discovered categories. We found that subjects show region dependent route preference according to their previous categorisation (Chi-Square test of Independence, X^2(1,19)= 13.34, p<.001). Therefore, place-names grouped into semantic categories can evoke a significant regionalisation. Also, the regionalisation effect can lead to different orientation behaviours even within the same environment according to the previously provided context.

Keywords: Spatial Cognition, Navigation, Virtual Reality

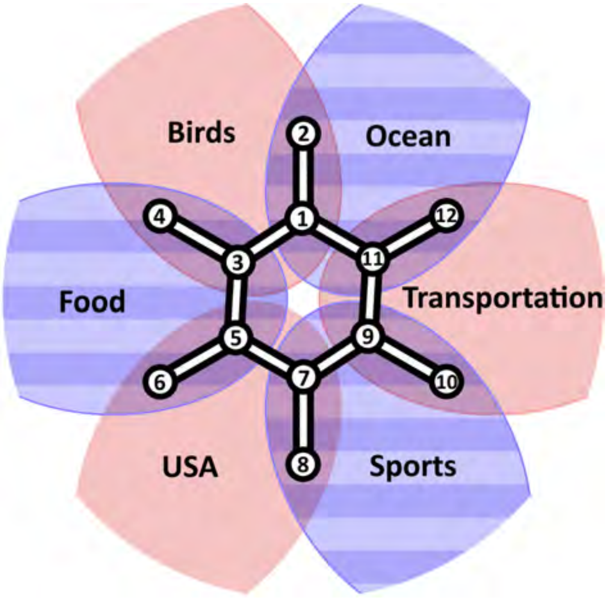


Figure 1: Each place had one place name, which could be included in either of two equally valid categorisations, respectively resulting in the red or blue regions. A classification task was used to bias the categories discovered by the subjects.

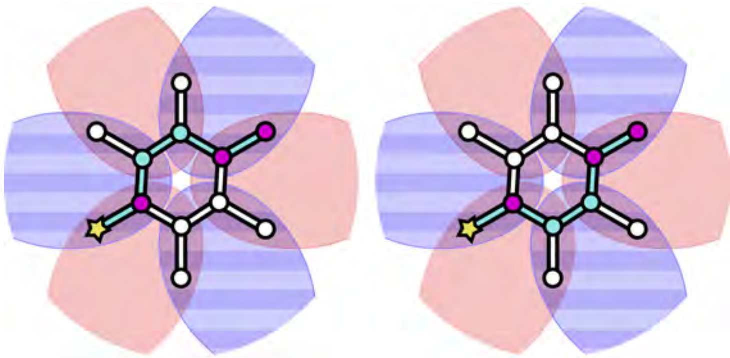


Figure 2: Possible shortest routes (i.e. without unnecessary u-turns) in a task with two equidistant solutions. When considering the Birds, USA and Transportation (red) as regions, the left path has two border crossings, while the right path only has one. Vice versa for Food, Ocean and Sports (blue) as regions. (Yellow Star = Start, Purple Circles= Destinations)

References

Burris, C. T., & Branscombe, N. R. (2005). Distorted distance estimation induced by a self-relevant national boundary. *Journal of Experimental Social Psychology*, *41*(3), 305–312.

Friedman, A., & Montello, D. R. (2006). Global-scale location and distance estimates: common representations and strategies in absolute and relative judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 333.

Schick, W., Halfmann, M., Hardiess, G., Hamm, F., & Mallot, H. A. (2019). Language cues in the formation of hierarchical representations of space. *Spatial Cognition & Computation*, *19*(3), 252–281.

Wiener, J. M., & Mallot, H. A. (2003). 'fine-to-coarse'route planning and navigation in regionalized environments. *Spatial cognition and computation*, *3*(4), 331–358.

A prototheory of climate cognition

Lucas Lörch
University of Mannheim

Being among the most severe threats to human civilization, climate change and the ecological crisis call for rapid collective action. However, our response to this crisis might be compromised by our cognitive system (Davis & Lewandowsky, 2021). The present work aims to develop a theory that describes how the unique features of our cognitive system might contribute to the emergence and maintenance of the ecological crisis.

Theory development in the present work is based on the *theory construction methodology* (TCM) proposed by Borsboom et al. (2021). This methodology comprises five steps, namely (1) identifying empirical phenomena that should be explained by the theory, (2) formulating a verbal prototheory that explains these phenomena, (3) translating the prototheory into a formal model, (4) checking the explanatory adequacy of the formal model, (5) evaluating the overall adequacy of the theory.

As a relevant phenomenon, I identified the finding that collective beliefs on climate change do not reflect the scientific consensus. For example, Leiserowitz et al. (2019) reported that although over 90 % of scientists agree that climate change is caused by human activity, 30 % of Americans belief it is due to natural causes. In a recent climate survey[1] by the European Investment Bank, only 47 % of Europeans and 39 % of Americans considered climate change to be a major threat to society. In the second step, I used abductive reasoning to developed a prototheory, i.e., a verbal theory that explains this phenomenon. In this prototheory, I assume that climate crisis (1) has political implications, (2) is subject of deliberate disinformation, (3) involves temporal and spatial offsets, and (4) is highly complex. Moreover, with reference to the cognitive system, I assume (a) motivated reasoning (Kunda, 1990), (b) a continued influence of disinformation (Lewandowsky & van der Linden, 2021), (c) an effect of psychological distance on construal level (Trope & Liberman, 2010), and (d) biased decision making in the face of complex information (Kahnemann & Tversky, 1974). I suggest that each of the named features of climate crisis can be combined with the respective feature of the cognitive system to explain the lacking societal consensus on the climate crisis and its solution. As a next step, I plan to translate the prototheory into a formal model.

In summary, the presented prototheory embeds well-established characteristics of the cognitive system into a single framework and thereby provides a comprehensive idea of the cognitive foundation of the emergence of the climate crisis. This demonstrates how "cognitive science can offer unique theoretical perspectives […] in the service of addressing global challenges." (Luo & Zhao, 2021, p. 22)

Keywords: theory development, climate change, climate crisis, cognitive biases, information processing

References

Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory Construction Methodology: A Practical Framework for Building Theories in Psychology. *Perspectives on Psychological Science*, 16(4), 756–766. https://doi.org/10.1177/1745691620969647

[1] https://www.weforum.org/agenda/2020/01/climate-change-perceptions-europe-china-us/

Davis C., J. & Lewandowsky, S. (2022). Thinking about climate change: look up and look around!. *Thinking & Reasoning*, https://doi.org/10.1080/13546783.2022.2041095

Kunda Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Leiserowitz, A., Maibach, E., Rosenthal, S., Kotcher, J., Bergquist, P., Ballew, M., Goldberg, M., & Gustafson, A. (2019). *Climate change in the American mind: November 2019*. Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.

Luo, Y., & Zhao, J. (2021). Attentional and perceptual biases of climate change. *Current Opinion in Behavioral Sciences*, 42, 22-26. https://doi.org/10.1016/j.cobeha.2021.02.010

Lewandowsky, S., & van der Linden, S. (2021). Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology*. https://doi.org/10.1080/10463283.2021.1876983

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463.

Tversky A., & Kahneman D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185:1124-1131.

Using emotional word ratings to extrapolated norms for valence, arousal, imagebility and concreteness: The German list of extrapolated affective norms (G-LEAN)

Jana Lüdtke & Katharina Gloria Hugentobler
Freie Universität Berlin

To study the influence of emotions on different levels of word, sentence and text comprehension, emotional lexicons describing the affective meaning of the presented words and text are a central building block. Word lists containing human rating data for theoretical well-established dimensions like valence and arousal are often limited in terms of coverage of the number of words they contain and the emotional variables they feature. Recent developments in the field of natural language processing and computational text analysis overcome these limitations as they allow the computation of automatically generated emotional lexicons (Westbury et al. 2015). One of the best established methods is an algorithm from Turney and Littman (2003) using distributional semantics (vector representations of words, VSM) to calculate emotional features for new words based on a vector space model and a label list containing prototypicaly words for the dimension of interest. The quality of the calculated norms lexicon hinges on the quality (Jacobs & Kinder, 2020) and length (Köper & Schulte im Walde, 2016) of the labels and the underlying VSM. Automatic generated label lists for the German language in current approaches often rely on human ratings taken from other languages and lack face validity, as they contain incomplete words, function words and less prototypical words. This hinders an application of these labels to other languages or other language models containing new German words.

We tested a new approach using German rating norms only and generated a label list which is independent of the vector space model used for the calculation of the new norms. The newly developed label lists for dimensions valence, arousal, imageability, and concreteness include only high frequent German words with high face validity (see Table one for examples). Using the algorithm by Turney and Littman (2003) we calculated a new list of extrapolated affective norms, the G-lean, and compared them with the available human ratings for the German language. Our approach produces results in line with state-of-the-art monolingual approaches to lexicon creation. Using our label lists produces norms with correlations with human ratings between .57 and .78 irrespective of the underlying VSM (see Table 2).

Keywords: affective norms, valence, arousal, automatic estimation, vector space model

Table 1
Examples for the label lists for four different dimensions

	dimension							
	Valence		*Arousal*		*Concreteness*		*Imagebility*	
	negativ	positive	Low	high	low	high	low	high
first three labels	Krieg Gewalt schlecht	Leben Liebe Zukunft	wenig Uhr Bereich	Angst Macht Tod	gut Zukunft Recht	Haus rund sprechen	sollen nie wollen	Uhr Mann Vater

Note: Each label list contains 60 words

Table 2
Correlations between norms for four different dimensions calculated with two different vector space models (VMS)

dimension[1]	VSM A[2]		VSM B[3]		Source for human ratings
	Pearson correlation[4]		Pearson correlation[4]		
	R^2	r	R^2	r	
Valence	.52	.72	.51	.72	extended version of BAWL (Conrad et al., 2014, Võ et al., 2009)
Arousal	.33	.57	.25	.50	extended version of BAWL (Conrad et al., 2014, Võ et al., 2009)
Imagebility	.38	.62	.40	.63	extended version of BAWL (Conrad et al., 2014, Võ et al., 2009)
Concreteness	.62	.79	.61	.78	concreteness merged table (Charbonnier & Wartena, 2020)

Notes:
[1] The values for the four dimensions were calculated based on 120 labels. The labels were generated separately for each dimension.
[2] For VSM A we complemented the SdeWac, a 0.88 billion words corpus derived from deWaC (https://wacky.sslmit.unibo.it/doku.php?id=corpora) by the German Poetry Corpus, a 16 million word corpus derived from the German text archive a(https://github.com/tnhaider/DLK).
[3] We used the German Wikipedia corpus with 3 million words for VMS B (https://fasttext.cc/docs/en/crawl-vectors.html)
[2,3] Both VSMs were calculated by using fastText with CBOW, position-weights, in dimension 300, with character n-grams of length 5, a window of size 5 and 10 negatives.
[4] For the calculation of the Pearson correlation for *Valence*, *Arousal* and *Imagebility* we could use human ratings for 5145 words to compare them with our new norms, for *Concreteness* we could use human ratings for 3741 words.

References

Charbonnier, J., & Wartena, C. (2020). Predicting the concreteness of German Words. In *SWISSTEXT & KONVENS 2020: Swiss Text Analytics Conference & Conference on Natural Language Processing 2020; Proceedings of the 5th Swiss Text Analytics Conference (SwissText) & 16th Conference on Natural Language Processing (KONVENS), CEUR Workshop Proceedings Vol. 2624*.

Conrad, M., Schmidtke, D., Võ, M. L.-H., & Jacobs, A. M. (2014). Emotion norms for German words: A large scale database of Valence, Arousal, and Imageability Rating values [Unpublished manuscript]. Experimental and Neurocognitive Psychology, Freie Universität Berlin.

Jacobs, A. M., & Kinder, A. (2020). Computing the Affective-Aesthetic Potential of Literary Texts. *AI.*, *1*(1), 11-27. doi: 10.3390/ai1010002

Köper, M., & Im Walde, S. S. (2016, May). Automatically generated affective norms of abstractness, arousal, imageability and valence for 350 000 german lemmas. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2595-2598).

Turney, P. D. and Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from as- sociation. *ACM Transactions on Information Systems*, *21*(4), 315–346. doi: 10.1145/944012.944013

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hof- mann, M., and Jacobs, A. (2009). The berlin affective word list reloaded (bawl-r). *Behavior Research Methods*, 41(2):534–538. doi: 10.3758/BRM.41.2.534

Westbury, C., Keith, J., Briesemeister, B. B., Hofmann, M. J., & Jacobs, A. M. (2015). Avoid violence, rioting, and outrage; approach celebration, delight, and strength: Using large text corpora to compute valence, arousal, and the basic emotions. *Quarterly Journal of Experimental Psychology*, *68*(8), 1599–1622. doi: 10.1080/17470218.2014.970204

Core hinges

Jakob Ohlhorst, University of Vienna

According to some developmental psychologists, animals have a stock of innate “core knowledge” or “core cognition” (Carey, 2009; Spelke & Kinzler, 2007). This core cognition is argued to be hard-wired and insulated from our ordinary cognition. That is, even when we come across contrary evidence, the core knowledge remains intact and is not rebutted or undermined. The idea is that core knowledge gives us a cognitive starter kit to come to grips with the environment. Some basic functions that core cognition modules fulfil are the recognition of conspecifics, causality, and relative quantities. A typical example would be the imprinting mechanisms that serve to designate something as the caretaker/parent for a young animal.

While the question of the innateness of knowledge played an important role in the epistemology of the Middle Ages and early Modernity, contemporary epistemologists have somewhat neglected the question, focussing more on acquired beliefs. The debate has instead been picked up by cognitive scientists (e.g. Carey, 2009) and philosophers of mind (Fodor, 1998). Meanwhile, I believe that contemporary epistemology can make a fruitful contribution to the issue. The fairly young field of *hinge epistemology* can deliver a framework to explain the epistemological functioning of core cognition and what role it plays for our other beliefs.

Hinge epistemology as a framework is inspired by Ludwig Wittgenstein’s *On Certainty* (1969). It focuses on *hinges*, fundamental convictions that are presupposed in our entire epistemic life. Metaphorically speaking, they are the unchanging background in front of which our epistemic life plays out. A typical hinge would be the conviction that other people have a mental life or that the world is older than a century. Epistemically speaking, hinges keep us from falling into a deep scepticism. Hinges are further characterised by the fact that they are not supported by evidence in any non-circular or non-regressive way; they are (psychologically) certain, and it would not make sense to doubt them.

I will argue for the thesis that core cognition consists of a special kind of innate hinges. Namely, I will argue that the cognitive mechanisms that constitute core cognition, e.g. Carey’s domain-specific learning mechanisms, encode hinge certainties – even though there is no proposition that is explicitly believed. This claim fits well with one of the most carefully worked out accounts of hinge epistemology: Danièle Moyal-Sharrock’s (2004) theory of hinges as being *animal*. Animal hinges are the pre-epistemic framework that enables us to act and live without having to continually re-examine the theoretical foundations of our practical and epistemic life. These animal hinges are non-propositional – they are neither truth-apt nor explicitly believed, but they manifest in our behaviour and ways of believing. Similarly, we take our core cognition simply for granted and do not scrutinise it.

Core knowledge differs from Moyal-Sharrock’s account insofar as core cognition does not play the same global role as she envisions. Instead, core cognition is our cognitive starting point to be able to successfully learn to navigate the world. For instance, many parts of our physical core knowledge are false and do not integrate into Newtonian, let alone relativistic or quantum, mechanics. Interestingly, these mistaken folk physical notions still remain active in adults, and they fulfil an important functional role because they are approximately true at the scales with which humans are confronted every day. Nevertheless, adults are not tethered to this core knowledge and they work around it, for instance when formalising physical events.

I will present our cognition of causality as an example to illustrate how core cognition aligns with our animal hinges:

(1) Michotte (1963) argues that we possess an innate perceptual module that decides if the movement of one object causes the movement of another object or not, depending on different criteria of spatial contact and temporal contiguity. According to Michotte, we consequently possess perceptual core knowledge of mechanical causation.

The animal hinge that is encoded by this module is a *rule* (*cf.* Coliva, 2015): Objects moving in a certain relation to each other *cause* each other's movements. This rule describes the cogniser's animal hinge certainty that there is such a thing as causality. But these hinge certainties are "reflex-like ways of acting and not propositions striking us immediately true" (Moyal-Sharrock, 2016, p. 32).

(2) Infants are able to categorise objects as either dispositionally inert or as dispositional agents capable of initiating movements. They ascribe different causal profiles to dispositional agents and inert objects respectively (Luo, Kaufman, & Baillargeon, 2009). Perception of Michottian mechanical causality as described above only applies to inert objects while dispositional agents are ascribed diverse causal capacities, including floating in mid-air. The hinge encoded here is the existence of agents; entities which are not subject to (bare) mechanical causality but to psychological causality instead.

The possibility of core knowledge's falsity and of its being overridden is incompatible with Moyal-Sharrock's account of animal hinges. These are fixed and immutable; a cogniser cannot work around them. Consequently, a core cognition account of hinges goes beyond the traditional Wittgensteinian accounts. An important criterion for whether core knowledge consists of hinges is whether it can be revised through evidence or not – hinges are essentially insensitive to evidential revision. Note that core cognition is argued to still be active in adults who, for instance, perceive Michottian causality. That is, the animal hinge is still active. *Vice versa*, introducing hinge epistemology to core cognition shows the fundamental role that these resources play for our very epistemic access to the world. Core cognition enables much more than just basic learning – sophisticated adult human cognition would not be possible without core cognition.

*Keywords***:** Concept nativism, Wittgenstein, core cognition, developmental psychology, hinge epistemology

References

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

Coliva, A. (2015). *Extended Rationality*. Basingstoke: Palgrave Macmillan.

Fodor, J. A. (1998). *Concepts: where cognitive sciences went wrong*. Oxford: Oxford University Press.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive Psychology*, *58*(4), 441–486. https://doi.org/10.1016/j.cogpsych.2008.11.001

Michotte, A. (1963). *The of Perception Causality*. New York: Basic Books.

Moyal-Sharrock, D. (2004). *Understanding Wittgenstein's On Certainty*. Basingstoke: Palgrave Macmillan.

Moyal-Sharrock, D. (2016). The Animal in Epistemology. In D. Moyal-Sharrock & A. Coliva (Eds.), *Hinge Epistemology* (pp. 24–47). Leiden: Brill.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96.

Wittgenstein, L. (1969). *Über Gewissheit - On Certainty*. (G. E. M. Anscombe & G. H. von Wright, Eds.). Oxford: Blackwell.

German-English Interlingual Homographs and Their Organization in the Bilingual Brain – Evidence From Semantic Similarity

Alina Palmetshofer[1], Carolin Dudschig[1], Fritz Günther[2] & Barbara Kaup[1]
Organisations: 1: Eberhard Karls University; 2: Humboldt-Universität zu Berlin

We investigated the organization of the bilingual lexicon by analysing how interlingual homographs – words like "GIFT" and "BLANK" that occur in more than one language and irrespective of their meaning (identical or not) have an identical spelling – are represented and comprehended by bilinguals. An innovative lexical decision task using only English-German homographs primes and targets besides pseudowords was implemented to investigate whether lexical decision (LD) time for the target word can be predicted by cosine similarity values that were obtained by vector space models. Cosine similarity was calculated for each prime-target pair and for their meanings in both languages. If L1 and L2 meaning are represented in a shared lexicon both L1 and L2 measures of semantic similarity should influence potential priming effects in a LD task. In contrast, if there is a clear separation between lexica, only the currently salient language should influence LD times. We recruited 45 German-English bilinguals in Germany who reported British English as their L1 and 45 German-English bilinguals with UK residency with German as their L1. An extensive language history questionnaire and vocabulary test LexTALE (Lemhöfer & Broersma, 2012) was administered at the end of the experiment to ensure participants' proficiency in both languages. Two experimental versions were created: One version contained only German text (e.g., task instruction) and a German cultural priming setup where stereotypical German scenes were shown in the beginning (e.g., pictures of the Brandenburger Tor) while in the other experimental version all information was given in English and typically British scenes were displayed as a cultural prime. This manipulation was introduced to make the language of the country of residency the salient language. All subjects were presented with the experimental version that matched their country of residency and asked to decide whether the presented items were real words without reference to any language. Thus, if the bilingual lexica are organized in a rather separate manner as hypothesized, and the language saliency manipulation works, predominantly the language of the country of residency should influence LD times. Contrary to expectations, model comparisons for the overall data set showed no significant interaction nor a main effect of cosine similarities or experimental version, suggesting that the manipulation did not work in this experimental setting. An unplanned post-hoc comparison revealed that for German participants (who were all living in the UK and received the matching British experimental version), German cosine similarity was a significant predictor ($F(1,159) = 4.55$, $p = .03$). However, no effect of cosine similarity was found for British participants. These results suggest that cosine similarity can be predictive for participants' LD time even when exclusively using interlingual homographs, but bilinguals may process homographs in their native language even if their L2 was made salient, which was the language spoken in participant's countries of residency. Based on those mixed findings, we set out to create a comprehensive data base of German-English interlingual homographs by selecting all possible candidates from large-scale corpora of the two languages. We envisage that this data base will be useful for future empirical studies such as the one described above. Our definition of interlingual homographs still includes words with identical meaning in both languages and was extended to foreign words in both languages. For the selected 1,653 words, we obtained ratings on age of acquisition, arousal, concreteness and valence by using best-worst scaling (e.g., Hollis & Westbury, 2018). Ratings were performed by monolinguals of German and English. We also included information about possible word classes and the words' phonological similarity between both languages. We derived pseudowords by means of intersecting the generated pseudo words for German and English by the *Wuggy* software (Keuleers & Brysbaert, 2010). Direct translation equivalents will be added to the data base to provide words with shared meaning but different form across languages that can then be compared with the interlingual homographs that share form, but not (necessarily) meaning. This data base will soon be available and allow researchers to pick suitable items for investigating the bilingual mental lexicon.

References

Hollis, G., & Westbury, C. (2018). When is best-worst best? A comparison of best-worst scaling, numeric estimation, and rating scales for collection of semantic norms. *Behavior research methods*, *50*(1), 115-133. https://doi.org/10.3758/s13428-017-1009-0

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods 42*(3), 627-633. https://doi.org/10.3758/BRM.42.3.627

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*, 325-343. https://doi.org/10.3758/s13428-011-0146-0

Discourse Illusions in L1 and L2

Clare Patterson
University of Cologne

Several types of linguistic illusion, such as semantic and grammatical illusions, have become well-studied phenomena in psycholinguistics. These illusions have in common that they contain violations which are systematically overlooked. As yet, illusions at the discourse level have not been discussed in this context. The aim of the current contribution is to establish whether discourse violations, similar to semantic and grammatical violations, are also overlooked in a systematic manner; this would constitute a discourse illusion.

(1) Laura is thinking about going out tonight. On the one hand, she feels like dancing, because a great DJ will be playing. On the other hand, she can sleep in tomorrow. She is finding it difficult to decide.

The discourse connective "on the one hand…on the other hand" sets up an expectation of contrast (Scholman et al., 2017). In (1), arguments in favour of Laura going out have been presented with "on the one hand". Contrasting arguments should thus be presented with "on the other hand". However, in (1), a further argument in favour is presented instead. Using an acceptability judgment paradigm, we test the extent to which discourse violations like (1) are overlooked. Acceptability judgments for discourse violations are expected to be highly variable in comparison to controls, similar to comparative illusions (Wellwood et al., 2018). Importantly, if judgments for discourse violations are significantly better than for other contradictory or nonsensical items, this will be taken as evidence that discourse violations are systematically overlooked. Furthermore, we test whether speakers are more susceptible to such illusions in their second language (L2) compared to their first language (L1), possibly due to differing processing strategies between L1 and L2. It is possible, however, that L2 speakers' difficulty in comprehending certain discourse connectives (e.g. Zuferrey et al., 2015) could be responsible for the effect.

Materials: Item acceptability was rated on a 1-7 scale. Items consist of 24 experimental items and 64 fillers. Two factors (Violation: violation; control, and Language: German; English) are manipulated to create four conditions for each experimental item (see Table 1). Each participant is tested in both their L1 (German) and their L2 (English), providing a within-participant comparison. Three types of discourse connectives are tested (8 items each) covering three broad categories: *contrast*, *similarity* and *denial of expectation*. A set of 16 'bad fillers' is included for comparison; these are either contradictory items (n=12) or are incoherent text passages (n=4) (see Table 2). The experiment consists of an English part (English items and fillers) and a German part (German items and fillers). The order of presentation of the two parts is alternated between participants. **Participants:** 251 students from a German university were tested; after exclusions 184 were included in the analysis, all German L1 speakers (self-reported). Participants' mean self-rated English comprehension was 5.6/7. Reading experience was measured using the Author Recognition Test for German (Grolig et al., 2020). **Results:** Data was analysed using Bayesian ordinal models. Effects are expressed as SD units. Overall, discourse illusions behaved like other illusion types in that control items were rated higher than violation items (0.26 [0.13, 0.38]) but bad fillers were rated lower than violations (-0.48 [-0.73, -0.22]) (Figure 1). Participants were more susceptible to illusions in their L2 compared to their L1, with German violations eliciting lower ratings than English violations (-0.42, [-0.51, -0.33]). The influence of reading experience was equivocal. **Outlook**: This large dataset offers the opportunity to explore several individual-level variables and make more precise effect-size predictions for future experiments. More generally, establishing the existence of discourse illusions opens the door to many questions about the underlying mechanisms supporting language processing at the discourse level and differences between L1 and L2 discourse processing.

Keywords: **linguistic illusions, discourse processing**, **L2**, **discourse connectives**

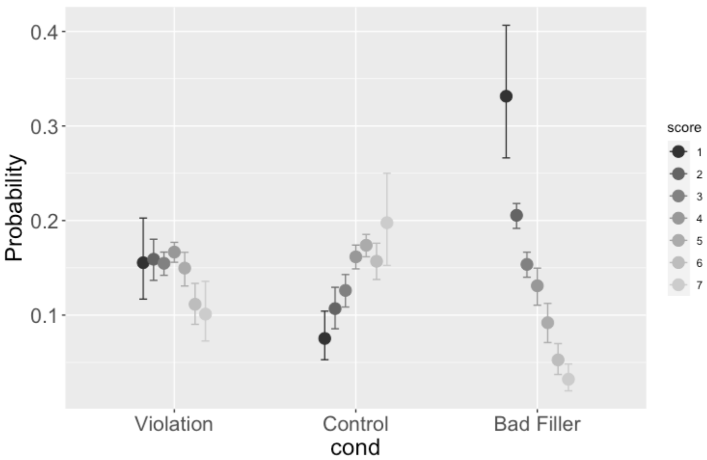
Table 1. Example experimental item in four conditions.

Violation, English	Laura is thinking about going out tonight. On the one hand, she feels like dancing, because a great DJ will be playing. Many of her friends are going to the party as well. On the other hand, she can sleep in tomorrow. She is finding it difficult to decide.
Control, English	Laura is thinking about going out tonight. On the one hand, she feels like dancing, because a great DJ will be playing. Many of her friends are going to the party as well. On the other hand, she can't sleep in tomorrow. She is finding it difficult to decide.
Violation, German	Laura überlegt, ob sie heute Abend feiern gehen soll. Einerseits hat sie Lust zu tanzen, weil ein guter DJ auflegt. Außerdem werden viele Freunde von ihr da sein. Andererseits kann sie morgen ausschlafen. Die Entscheidung fällt ihr heute sehr schwer.
Control, German	Laura überlegt, ob sie heute Abend feiern gehen soll. Einerseits hat sie Lust zu tanzen, weil ein guter DJ auflegt. Außerdem werden viele Freunde von ihr da sein. Andererseits kann sie morgen nicht ausschlafen. Die Entscheidung fällt ihr heute sehr schwer.

Table 2. Example of bad filler items.

Contradictory filler	After her car accident, Clara had to go to court. She feared the judge, because he was so patient. In the end, Clara had to pay a large fine.
Incoherent text	Andrew is thinking about becoming a caricaturist, but he knows that art is not an easy way to making a living. He gave his uncle a piece of chewing gum, in case the car took too long to arrive.

Figure 1. Probabilities for each score (1-7) for the violation, control and bad filler items.



References

Grolig, L., Tiffin-Richards, S. P., & Schroeder, S. (2020). Print exposure across the reading life span. *Reading and Writing*, *33*, 1423–1441.

Scholman, M., Rohde, H., & Demberg, V. (2017). "On the one hand" as a cue to anticipate upcoming discourse structure. *Journal of Memory and Language*, *97*, 47-60.

Wellwood, A., Pancheva, R., Hacquard, V., & Phillips, C. (2018). The anatomy of a comparative illusion. *Journal of Semantics, 35,* 543–583.

Zuferrey, S., Mak, W., Degand, L., & Sanders, T. (2015). Advanced learners' comprehension of discourse connectives: the role of L1 transfer across on-line and off-line tasks. *Second Language Research*, *31*, 389-411.

Using Cognitive Models and EEG Data To Investigate Spatial Cognition

Kai Preuss (preuss@tu-berlin.de) and Nele Russwinkel (nele.russwinkel@tu-berlin.de)
Kognitive Modellierung dynamischer Mensch-Maschine-Systeme
Institut für Psychologie und Arbeitswissenschaft, Technische Universität Berlin

In the field of spatial cognition, spatial transformation stands as a barely researched topic. Yet it is this cognitive ability that enables us to identify, plan and decide on possible interactions with an object, proving paramount for daily life. Classic experimental paradigms of mental (spatial) transformation include mentally rotating abstract objects (*mental rotation,* Shepard & Metzler, 1971) or mentally folding cube patterns (*mental folding,* Shepard & Feng, 1972) to differentiate between two presented objects. Harris et al. (2013) identified similarities and differences between mental rotation and mental folding formalized as *rigid* and *non-rigid* transformations. The *rigidity* of mental spatial transformations may therefore reflect itself in discernible cognitive processes, but this has proven hard to differentiate by behavioural studies alone – reaction times or error rates by themselves do not suffice to discern underlying cognitive mechanisms.

Cognitive modelling offers unique insight into mental activity during task-solving by simulating an experiment participant's process of reasoning. The cognitive architecture ACT-R in particular is based on a modular structure, with distinct modules for specific mental functions working in parallel (Anderson & Lebiere, 2014). To allow for realistic processing of spatial information, an additional module was theorized and implemented into the architecture, representing a dedicated cognitive mechanism for spatial cognition (Preuss et al., 2019). At the same time, mental rotation and folding experiments were conducted with EEG recording, and cognitive models mirroring participants in these experiments have been developed with ACT-R (Preuss et al., 2019; Preuss & Russwinkel, 2021). Subsequently, the module activity of each cognitive model was compared to the behavioural data of its respective experiment to shine light on involved brain areas and intra-trial processing phases.

For the mental rotation study, predictions by the cognitive model were compared to Hidden semi-Markov Models (*HsMM*, Borst & Anderson, 2015) resulting from post-processing the experiment's EEG data. HsMM-EEG finds phases of common EEG activity in experiment trials, thereby suggesting so called cognitive processing stages apparent during task solving. These stages are however devoid of semantic information, i.e. their individual meaning or function. Module activity predictions by the ACT-R model were able to not only offer functional explanations on the HsMM stages, but on possible strategy choices as well. For both the mental rotation and mental folding study, cognitive model output was matched with an Independent Component Analysis (*ICA*) applied to this experiment's EEG recordings (Hilton et al., 2022). This identified clusters of brain activity maximally distinct from one another in regard to their activation during a trial across different experiment conditions. From these clusters, brain areas with a commonly ascribed function were chosen as candidates for comparison with their respective ACT-R module counterpart. Subsequent analysis found module activity simulated by the cognitive model to align with actual brain activity during the experiment. Notably, close matches between model and data were achieved with visual, memory and image-building processes.

Both approaches help understand and differentiate mental spatial transformations while serving as proof-of-concept for the viability of advanced applications of cognitive models in conjunction with high-dimensional data.

Keywords: spatial cognition, mental spatial transformation, ACT-R, EEG

Figures

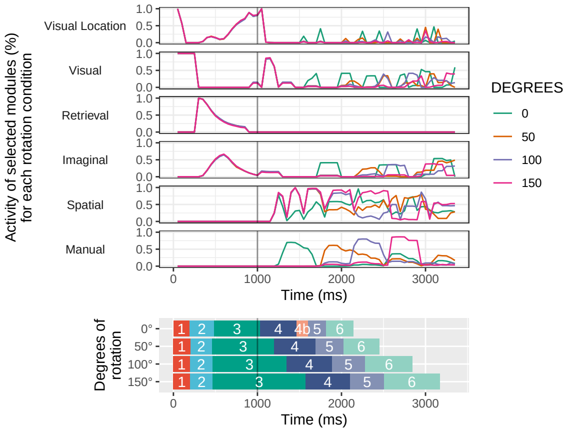


Figure 1: Comparison of selected module activity (*top*) and HsMM stages (*bottom*) over the course of a mental rotation trial, separated by rotation condition.

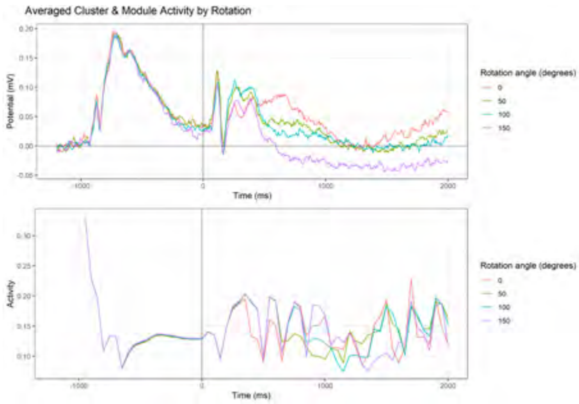


Figure 2: Comparison of aggregated EEG cluster activity (*top*) and module activity (*bottom*) over the course of a mental rotation trial, separated by rotation condition.

References

Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought*. Psychology Press.

Borst, J. P., & Anderson, J. R. (2015). The discovery of processing stages: Analyzing EEG data with hidden semi-markov models. *NeuroImage*, *108*, 60–73.

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: similarities and differences between mental rotation and mental folding. *Cognitive Processing*, *14*(2), 105–115.

Hilton, C., Raddatz, L., & Gramann, K. (2022). A general spatial transformation process? Assessing the neurophysiological evidence on the similarity of mental rotation and folding. *Neuroimage: Reports*, *2*(2).

Preuss, K., Raddatz, L., & Russwinkel, N. (2019). An implementation of universal spatial transformative cognition in ACT-R. In T. D. Stewart (Ed.), *Proceedings of the 17th international conference on cognitive modelling* (pp. 144–150). Waterloo, Canada: University of Waterloo.

Preuss, K., & Russwinkel, N. (2021). Cognitive modelling of a mental rotation task using a generalized spatial framework. In T. D. Stewart (Ed.), *Proceedings of the 19th international conference on cognitive modelling* (pp. 220–226). University Park, PA.

Shepard, R., & Feng, C. (1972). A chronometric study of mental paper folding. *Cognitive Psychology*, *3*(2), 228–243.

Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Towards understanding spatial illusions in architecture - a pilot study exploring factors influencing illusive perspectives

Vojtěch Rada[1], Julia Frankenstein[2], David Sedláček[1], Zdeněk Míkovec[1], and Constantin A. Rothkopf[2]

[1]Department of Computer Graphics and Interaction, Czech Technical University in Prague
[2]Psychology of Information Processing, Technical University of Darmstadt

Abstract

Illusive perspectives (forced perspectives) have been used in art and architecture to create certain space perceptions (e.g., let built space appear larger or deeper), space illusions (e.g., let persons or objects in certain positions appear larger or distorted) and artificial effects (e.g., nudge the visitor to stand at a certain position, i.e., forced perspective). To induce these spatial visual illusions, built spaces like hallways have been constructed such that visual cues induce increased depth if perceived by the visitor from a certain position (see Figs. 1a, b). [Dunning (1991)]

In our study, our aim was to determine how the properties of a distorted hallway (e.g., angle of the walls, amount of distortion) influence the perceived illusory space. We created a simplified VR model of an illusive hallway (see Figs. 2a, 2b) inspired by a historical example designed by Italian architect Francesco Borromini, and conducted a qualitative pilot study.

Participants wearing a head-mounted display (HMD) experienced ten virtual hallways with different properties. They were placed in the central point in front of a hallway (see Figs. 2a, 2b) and were asked to use a joystick to virtually move in front of and inside the virtual hallway. Hallway properties included geometrical factors (e.g., angle of hallway walls, hallway length and inclination of the hallway floor, see Figs. 2c, 2d), appearance of the hallway and lighting conditions (e.g., smoothness of edges, fog). Participants were asked to indicate whether a specific corridor was perceived to extend in depth or whether it could be detected as an illusory corridor. The interviews revealed that virtual movement (especially changing the user's vantage points towards the sides) and hallway geometry (the angle of hallway walls) were considered to be the most important factors for the illusion.

In our ongoing work, we are currently designing an experiment based on psychophysical methods to determine how the factors of hallway geometry (e.g., angle, amount of distortion, depth) identified in this pilot study interact to create the respective space illusion, and to what extent the size of the area in which the illusion can be experienced by the visitor changes based on hallway geometry factors.

Keywords: Forced perspective, spatial cognition, architecture, Ames room, depth perception

Figure 1: Santa Maria presso San Satiro by Donato Bramante 1a) view from the entrance, where illusion works, 1b) view from side, where the illusion breaks, Source: Personal archive

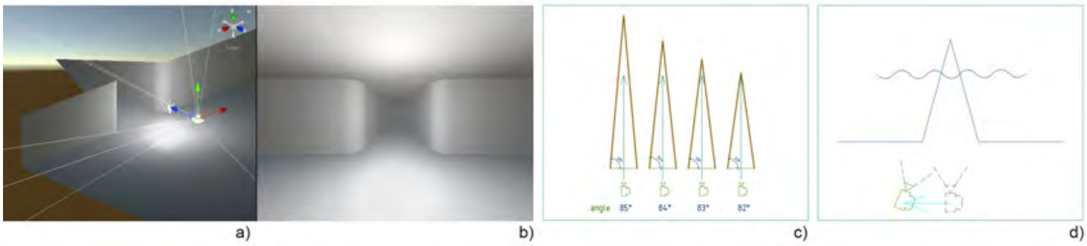


Figure 2: 2a,b) Simplified VR representation of an illusive hallway (ortographic and perspective view) 2c) Example of different angles between perceived and actual walls of the hallway 2d) Virtual movement of the participants and distance fog (wave line), Source: Personal archive

References

Dorward, F. M., & Day, R. H. (1997). Loss of 3-d shape constancy in interior spaces: The basis of the ames-room illusion. *Perception*, *26*(6), 707–718.

Dunning, W. V. (1991). *Changing images of pictorial space: a history of spatial illusion in painting*. Syracuse University Press.

Dwyer, J., Ashton, R., & Broerse, J. (1990). Emmert's law in the ames room. *Perception*, *19*(1), 35–41.

Glennerster, A., Gilson, S., Tcheang, L., & Parker, A. (2010, 10). Perception of size in a 'dynamic ames room'. *Journal of Vision - J VISION*, *3*, 490-490. doi: 10.1167/3.9.490

Gogel, W. C., & Mershon, D. H. (1968). The perception of size in a distorted room. *Perception & Psychophysics*, *4*(1), 26–28.

Hosseini, S. V., Alim, U. R., Oehlberg, L., & Taron, J. M. (2021). Optically illusive architecture (oia): Introduction and evaluation using virtual reality. *International Journal of Architectural Computing*, *19*(3), 291–314.

Peer, A., & Ponto, K. (2019). Mitigating incorrect perception of distance in virtual reality through personalized rendering manipulation. In *2019 ieee conference on virtual reality and 3d user interfaces (vr)* (pp. 244–250).

The perception of Spanish vowels in adverse listening conditions: An investigation of adults with typical hearing, children with typical hearing, and children with hearing impairment

Marcel Schlechtweg & Mark A. Gibson
Carl von Ossietzky Universität Oldenburg, Germany; Universidad de Navarra, Spain

Accessing pieces stored in the mental lexicon represents a complex phenomenon, and is especially challenging (a) if the individual cannot hear items in the way peers do and (b) if this process happens in a realistic environment with adverse listening conditions. Hearing impairment has been shown to affect several facets of language, such as vowels, consonants, or stress (see, e.g., Davies et al. 2020; Frank et al. 1987; Gilbert & Pisoni 2012). We ask here (a) how well Spanish adults with typical hearing (ATH), Spanish children with typical hearing (CTH), and Spanish children with hearing impairment (CHI) access the five vowel phonemes of their native language in noise / adverse listening conditions, (b) which vowels are correctly identified more and which ones less frequently, and (c) if the groups behave in a similar way with respect to the vowels they detect easily / with difficulty. We refer to data from similar, but not identical, experiments, one testing the two groups of children and one testing the adults.

We have so far tested 5 ATH (3 female, 2 male, 18–35 years), 4 CTH (4 male, 6, 8, 8, and 10 years), and 6 CHI (4 male, 8, 9, 10, and 13 years; 2 female, 10 and 12 years).[1] They heard the isolated syllables /da, de, di, do, du/, produced by a male and a female speaker, and were requested to select the syllable they heard by clicking on one of the five textboxes they saw on the screen. The syllables were embedded in two types of noise, background babble and signal-to-noise ratio. Note here that the experiment for the children (ExpCh) was different from the experiment for the adults (ExpAd). In ExpCh, background babble means that there were always 6 speakers; in ExpAd, there were 1, 2, 3, 4, 6, 8, 10, or 12 speakers. Also, in ExpCh, there was always more signal than noise, while, in ExpAd, there was either more noise or a relation of 1 to 1 of signal and noise. We used a different experiment for the children in order to keep the task manageable, especially for the CHI.

We observe in Table 1 that, first, CTH responded overall more accurately than CHI, second, /o/ and /u/ are harder to detect than the other vowels in all of the three groups, and, third, CHI face immense difficulties in perceiving /u/ and still non-negligible difficulties in perceiving /o/ (and /i/).[2] We intend to collect more data, and interpret these preliminary results as first evidence for the general difficulty in accessing phonemes related to two acoustically close vowels (/o/ and /u/), which are similar in terms of F2 (both are back vowels) and F3 (both produced with rounded lips), in noise (see also, e.g., Bradlow 1995; Reetz & Jongman 2009). F1 (vowel height) could in principle distinguish between the two, but does so only to a limited extent and seems to be dominated by F2 and F3, especially in the absence of any visual input, which would possibly show the difference between the two vowels (lip aperture / jaw angle). Our results are interpreted against the background of the role hearing impairment and noise play in the process of accessing phonological categories.

Table 1. Mean accuracy by group and vowel (in percent)

Vowel	ATH	CTH	CHI
a	77	96	95
e	76	98	95
i	73	97	70
o	55	72	63
u	51	88	48
Total	**66**	**90**	**74**

[1] The data from one CTH was discarded due to a very low general response accuracy (29 percent).
[2] Note that we cannot directly compare the two groups of children to the group of adults since they were tested on different experiments, which explains why the adults show lower accuracy rates for most cases.

Keywords: Hearing impairment, noise, children, vowels, Spanish

References

Bradlow, A. R. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America*, 97(3), 1916-1924.

Davies, B., Rattanasone, N. X., Davis, A., & Demuth, K. (2020). The acquisition of productive plural morphology by children with hearing loss. *Journal of Speech, Language, and Hearing Research*, 63, 552-568.

Frank, Y., Bergman, M., & Tobin, Y. (1987). Stress and intonation in the speech of hearing-impaired Hebrew-speaking children. *Language and Speech*, 30(4), 339-356.

Gilbert, J. L., & Pisoni, D. B. (2012). Vowel perception in listeners with cochlear implants. In Ball, M. J., & Gibbon, F. E. (eds.), *Handbook of vowels and vowel disorders* (pp. 386-405). Psychology Press.

Reetz, H., & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics, and perception*. Wiley-Blackwell.

Investigating the effects of facial expressions and color cues on processing negated and affirmative sentences

Emanuel Schütt[1], Merle Weicker[2], Carolin Dudschig[1]
[1] Department of Psychology, Language and Cognition Research Group, University of Tübingen
[2] Department of Psycholinguistics and Didactics of German, Goethe University Frankfurt/Main

It is assumed that humans make use of linguistic information (e.g., semantic knowledge about words) and non-linguistic information (e.g., speaker identity information) during language comprehension. Previous research has particularly shown that non-linguistic information such as world knowledge (e.g., Dudschig et al., 2016; Hagoort et al., 2004) and characteristics of a speaker (e.g., Rück et al., 2017; Van Berkum et al., 2008) is integrated rapidly and often in parallel to linguistic information when sentences are processed. Interestingly, specific facial expressions (i.e., the "not face") and colors (green vs. red) have been identified as non-linguistic markers of negation processing (Benitez-Quiroz et al., 2016; Dudschig et al., in press). In three pre-registered experiments, we investigated whether these markers are able to influence the comprehension of sentential negation and affirmation.

Recently, it has been suggested that producing and comprehending negated utterances is associated with the so-called "not face" – a unique and universal negation marker combining facial signs of contempt, anger, and disgust (Benitez-Quiroz et al., 2016). We explored whether perceiving non-linguistic information in terms of the "not face" can facilitate sentential negation processing. Participants read affirmative and negated sentences (e.g., "Ja, ich will ein Glas Wasser haben" ["Yes, I want to have a glass of water"] vs. "Nein, ich möchte nicht ins Kino gehen" ["No, I do not like to go to the movies"] in a self-paced manner (Experiment 1) or judged the sensibility of affirmative and negated sentences (Experiment 2). Prior to the onset of each sentence, a "not face" or a positive control face was displayed. Traditional *p* values as well as Bayes factors from linear mixed effects analyses indicated that there was no effect of the facial expressions on comprehension. In particular, neither reading times (Experiment 1; p = .134, BF = 0.09) nor sensibility judgment times (Experiment 2; p = .100, BF = 0.12) for negated sentences were significantly faster when these were preceded by a "not face". Thus, the "not face" appears not to affect sentential negation processing.

In the third experiment, we investigated whether green and red color cues influence the processing of affirmative and negated sentences. In a recent study, Dudschig et al. (in press) showed that performing "yes" ("no") responses in a lexical decision task is facilitated when the color of the corresponding response button is green (red). We aimed to examine whether this association can be extended to a sentence comprehension task, with green color cues promoting the comprehension of affirmative sentences and red color cues promoting the comprehension of negated sentences. Participants again judged the sensibility of affirmative and negated sentences. Green and red color cues were presented prior to the onset of each sentence. Similar to Experiments 1 and 2, we did not observe an effect of the non-linguistic information on comprehension: There were no processing advantages when color cues matched the sentence polarity (affirmation-green; negation-red), p = .701, BF = 0.03.

In sum, our results did not reveal any influence of the "not face" or green and red color information on processing negated and affirmative sentences. This might suggest that this type of non-linguistic information was not directly integrated during sentence comprehension. We will discuss potential reasons for our findings, including that the specific non-linguistic cues under investigation are not crucial for sentence comprehension; participants did manage to ignore the non-linguistic cues as they were task-irrelevant; negation and affirmation could be higher-level abstract operations that work independently from any non-linguistic influences.

Keywords: language comprehension, pragmatics, negation, affirmation, not face

References

Benitez-Quiroz, C. F., Wilbur, R. B., & Martinez, A. M. (2016). The not face: A grammaticalization of facial expressions of emotion. *Cognition*, *150*, 77-84. https://doi.org/10.1016/j.cognition.2016.02.004

Dudschig, C., Kaup, B., & Mackenzie, I. G. (in press). The grounding of logical operations: The role of color, shape, and emotional faces for “yes” or “no” decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Dudschig, C., Maienborn, C., & Kaup, B. (2016). Is there a difference between stripy journeys and stripy ladybirds? The N400 response to semantic and world-knowledge violations during sentence processing. *Brain and Cognition*, *103*, 38-49. https://doi.org/10.1016/j.bandc.2016.01.001

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438-441. https://doi.org/10.1126/science.1095455

Rück, F., de la Vega, I., Leuthold, H., & Kaup, B. (2017). Integration of visual information about the speaker during sentence processing. In S. Featherston, R. Hörnig, R. Steinberg, B. Umbreit, & J. Wallis (Eds.), *Proceedings of Linguistic Evidence 2016: Empirical, Theoretical, and Computational Perspectives*. University of Tübingen. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77659

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580-591. https://doi.org/10.1162/jocn.2008.20054

Why do alerting signals increase congruency effects in the flanker task? An examination of boundary conditions

Verena C. Seibold, University of Tübingen, Germany

Various studies have shown that the presentation of an uninformative alerting signal (AS) increases congruency effects in response conflict tasks such as the flanker task (e.g., Fischer et al., 2012). Yet, the mechanisms that underly this alerting x congruency interaction are still debated (for an overview see, e.g., Seibold, 2018). In the present study, I aimed at further elucidating this matter by exploring two potential boundary conditions: First, I investigated whether the emergence and / or the size of the alerting x congruency interaction depend(s) on the type of AS. Second, I probed the view that the alerting x congruency interaction depends on stimuli that have pre-existing stimulus-response directional associations such as arrows or numbers (see Kahan & Zhang, 2019). According to this view, alerting specifically enhances the activation of these stimulus-response directional associations (e.g., the activation of a left- or right-hand response by a left- or right-pointing arrow) and thereby leads to a larger congruency effect. To put this view to an empirical test, I investigated whether the alerting x congruency interaction can also be observed for stimuli that do not have these pre-existing stimulus-response directional associations (i.e., letters; see also Seibold, 2018).

The experiment was conducted online. Participants (N = 36) performed a letter flanker task, in which they responded to a central target (the letter “a” or “b”) by means of a left- or right-hand button press. The target (e.g., “a”) was surrounded by four flankers that were associated either with the same response (e.g., “A”; *congruent*) or with the opposite response (e.g., “B”; *incongruent*; see Figure 1). In half of the trials, an AS was presented shortly before the task display. The AS could be *visual* (i.e., asterisks above and below fixation), *auditory* (i.e., a 1,000 Hz sine tone), or *audio-visual* (i.e., a combination of the visual and auditory AS). Overall, participants completed 384 trials, which were subdivided into eight blocks with 48 trials each and were preceded by one practice block. The independent variables *Congruency* (congruent; incongruent) and *Alerting Condition* (without AS; visual; auditory; audio-visual) were varied randomly within blocks. The stimulus-response assignment was counterbalanced across participants. Reaction time (RT) and error rate were measured as dependent variables.

Repeated-measures ANOVAs on mean RT and mean error rate with the factors Congruency and Alerting Condition revealed longer RT, $F(1, 35) = 259.44$, $p < .001$, $\eta_p^2 = .88$, and more errors, $F(1, 35) = 39.36$, $p < .001$, $\eta_p^2 = .53$, in the incongruent condition compared to the congruent one as well as a main effect of Alerting Condition on RT, $F(3, 105) = 61.24$, $p < .001$, $\eta_p^2 = .64$. Furthermore, an alerting x congruency interaction was present in both RT, $F(3, 105) = 4.36$, $p = .006$, $\eta_p^2 = .11$ (see Figure 2, left panel), and error rate, $F(3, 105) = 9.58$, $p < .001$, $\eta_p^2 = .21$ (see Figure 2, right panel). Planned comparisons showed that, for RT, the congruency effect increased with a visual AS compared to the condition without AS ($p = .009$), while there was no further increase from the visual to the auditory to the audio-visual AS (ps > .11). However, an AS type-dependent increase of the congruency effect was present in the error rate: The audio-visual AS increased the congruency effect more than the auditory AS ($p < .001$); there was no difference between the auditory and the visual AS ($p > .32$).

In line with previous studies, a sizeable alerting x congruency interaction was observed. Even though the size of this interaction did not depend on the type of AS with respect to response speed (RT), it did so with respect to response accuracy (error rate). This result indicates that physical properties of the AS may indeed play a role for the size of the alerting x congruency interaction and should be considered in the future when comparing results across studies. Furthermore, the observation that alerting increased the congruency effect for letters in the present study questions the view that the alerting x congruency interaction depends on pre-existing stimulus-response directional associations. Rather, it may depend on the potential of a stimulus to form (sufficiently strong) stimulus-response associations.

*Keywords***:** Alerting, Response Conflict, Flanker Task

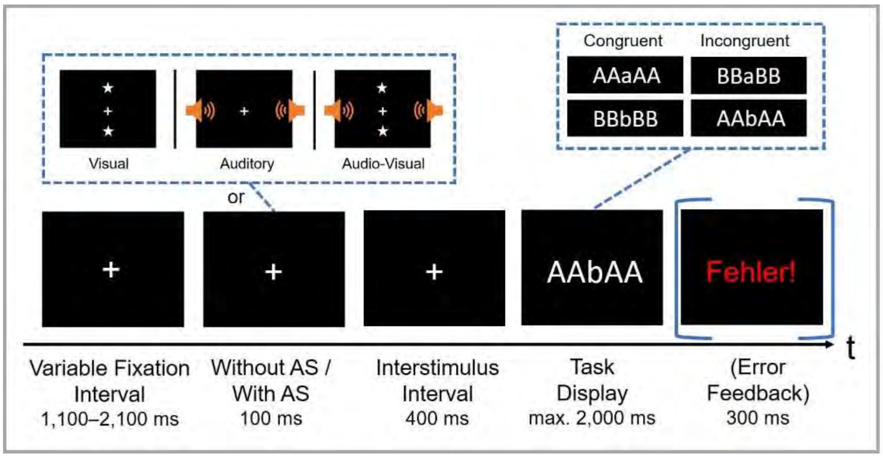


Figure 1. Illustration of the trial procedure. Participants responded to a central target letter (a or b) surrounded by (congruent or incongruent) flanker letters (A or B). In half of the trials, either a visual or an auditory or an audio-visual alerting signal (AS) was presented before the task display. Error feedback ("Fehler!", "Zu langsam!") was provided in case of an incorrect or a too slow response.

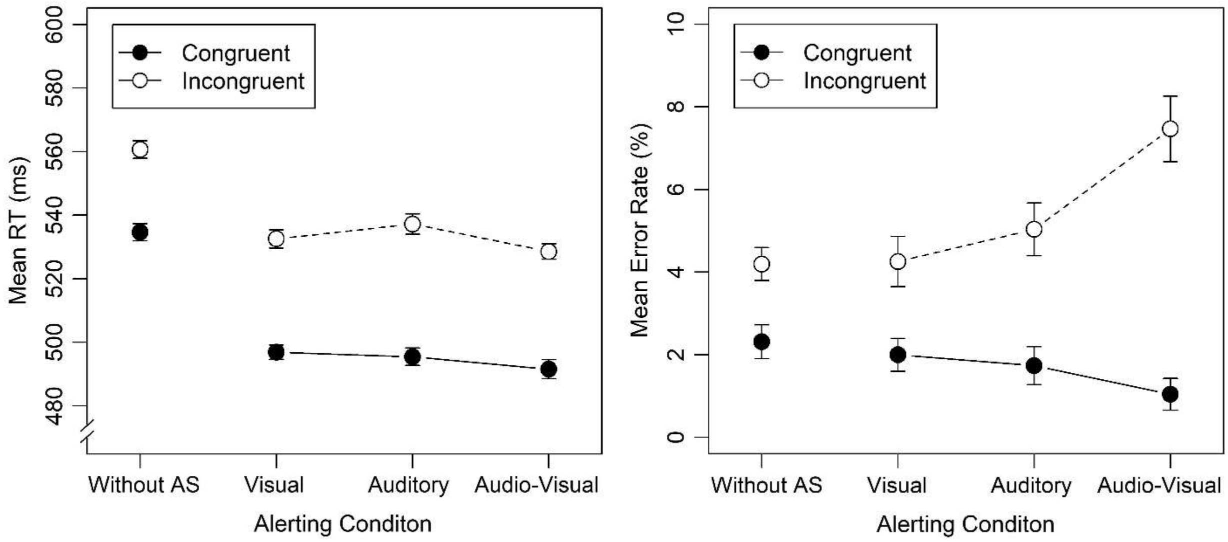


Figure 2. Mean reaction time (RT; left panel) and mean error rate (right panel) as a function of Congruency (congruent; incongruent) and Alerting Condition (without AS; visual; auditory; audio-visual). Error bars denote +/- 1 *SEM*, being corrected for within-subjects designs according to Morey (2008).

References

Fischer, R., Plessow, F., & Kiesel, A. (2012). The effects of alerting signals in action control: Activation of S–R associations or inhibition of executive control processes? *Psychological Research*, *76*, 317–328. https://doi.org/10.1007/s00426-011-0350-7

Kahan, T. A., & Zhang, H. (2019). Ready to be distracted: Further evidence that the alerting-congruency interaction requires stimulus-response directional associations. *Visual Cognition*, *27*, 760–767. https://doi.org/10.1080/13506285.2019.1680586

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Seibold, V. C. (2018). Do alerting signals increase the size of the attentional focus? *Attention, Perception, & Psychophysics*, *80*, 402–425. https://doi.org/10.3758/s13414-017-1451-1

Does contrastive attention guidance facilitate action recall? - An eye-tracking study.

Amit Singh, Katharina J. Rohlfing
Paderborn University, Germany

Teaming up with AI systems, humans could be guided for action performance. However, little is known about what kind of guidance is facilitating human performance. Studies in social science suggest that when humans use explanations, they are often contrastive (Miller, 2021) Contrastive explanations promote easier communication, fine grained understanding, and reduce cognitive load (Lipton, 1990). Although there are many studies investigating the linguistic representational aspects of contrastive explanation there are very few investigating it in a domain of motion event understanding. The following study fills this gap by bringing together contrastive explanation and studies on motion event. A motion event consists of a source, a figure, and a goal (Talmy, 2000), which provides a possibility to investigate the effect of different forms of verbal guidance on the action understanding. We reasoned that if contrastive verbal explanation is easier to understand generally, it might also lead to better understanding of action sequences. Previous research in psycholinguistics suggests a goal prominence in motion event, whereby participants retain a robust memory of goal than the source or the path (Papafragou, 2010). Informed by the research background, in this study we investigated whether, a) contrastive verbal guidance facilitates the motion event recall as in line with the contrastive explanation literature, and b) to what extent the goal prominence is affected by such guidance.
Stimulus and Method: Participants were presented video stimulus (N = 20) in which a ball was moved in relation to three landmark objects (Fig.1). Crucially, each action sequence was performed with a verbal guidance illustrating the path of the motion. The guidance was designed along two variables accounting for the contrastiveness in terms of verbal utterance (assertion or negation) and performed movements (i.e., motion path). For the verbal utterances, we used an Assertive(A) in contrast to Negative(N) verbal instructions. We chose negation for the reasons that a) negation has been shown to activate the alternate representation (Kaup, 2006), which can be contrasted against its positive counterpart, and b) negation guides, what not to do in addition to what has to be done and might reduce the goal bias which is prevalent otherwise (e.g., not down/up). For the performed movements, the ball followed either a contrastive (C) e.g., (up-down or down-up) vs non-contrastive (NC) (up-up or down-down) configurations. The recall was immediately assessed after each trial where participants turned 180 deg. from the eye-tracker and performed motion sequence without time constraint. To measure the goal bias, we calculated the fixation proportion on the final object path (alternate path) in the late window timeline (fig. 1), when the ball moved in the opposite direction (actual path).
Results: The results (N = 29) show the effect of contrast and verbal guidance on event recall and goal bias (fixation). To visualise activation of goal; the mean fixation on alternate Area of Interest (AOI) was calculated (fig.1 c(ii), nc(i)). The recall was coded 0 for incorrect or 1 for correct responses in pre- and post-windows for each correct path performed. A mixed effect logistic model was fitted with voice and path condition as fixed effect and random intercept for subjects. Figure 2(b) shows the recall for C and NC conditions yielding a main effect of path, such that recall for C>NC. Taking NV (No Voice) as the baseline, there was an overall better recall for AA voice condition in C and NC path conditions which was not surprising. Participants could better recall the paths when it was accompanied by assertive voice guidance. Crucially there was a significant interaction between path and voice such that for NA voice condition the recall was higher in C than the NC path condition. As predicted, we find that contrastive verbal guidance is facilitative when combined with a contrastive path and otherwise for NC the verbal contrast was detrimental. Moreover, the recall for NC path was higher in NN voice condition which might be due to the repetition of negation which was absent in C paths. For goal activation, we performed GCA analysis (Mirman et al., 2008) with 3rd order polynomial on logit transformed fixation data with voice as fixed effect. Fig. 2(a) shows the fixation pattern to the goal object. There was a main effect of voice, such that fixation for NV>AA>NN>NA>AN, suggesting that negation significantly reduced the fixation on alternate goal path during post window motion which was otherwise maximum when there was no guidance in NV baseline.

***Keywords**:* Contrastive Explanations, Action Recall, Motion Event, Negation, Attention.

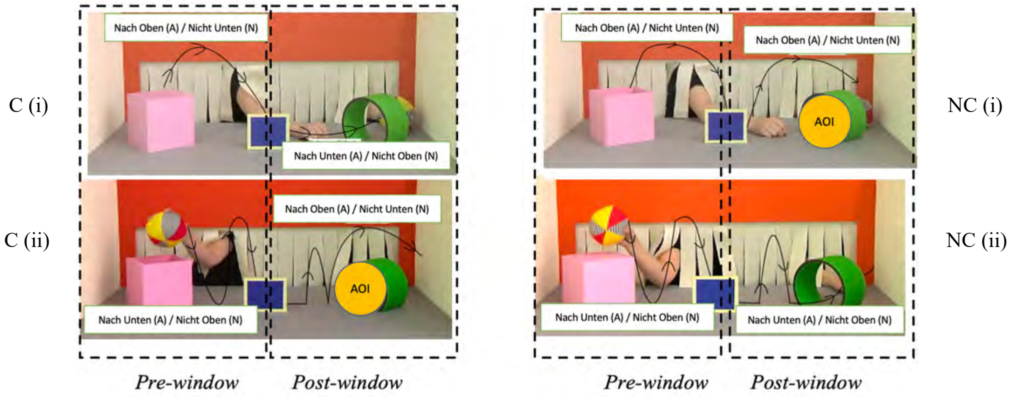


Figure 1. Path conditions, C, NC, C, NC (from top left to bottom right). Each movement window was preceded by either an Assertive(A) e.g., “nach oben” [towards up] or Negative(N) e.g., “nicht oben” [not up] voice predicate to create different degrees of contrasts including No Voice (NV) as baseline (AA, NA, AN, NN, NV).

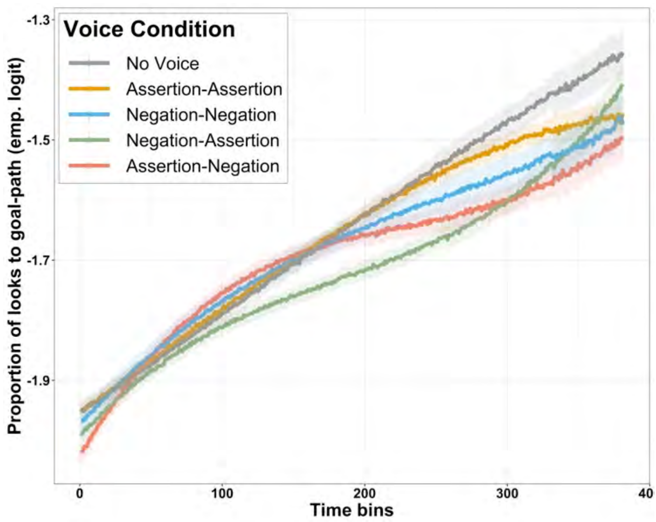


Figure 2 (a). Mean proportion of fixation on goal alternate path. Results are shown for analysis time-window of post-window region highlighted in fig 1.

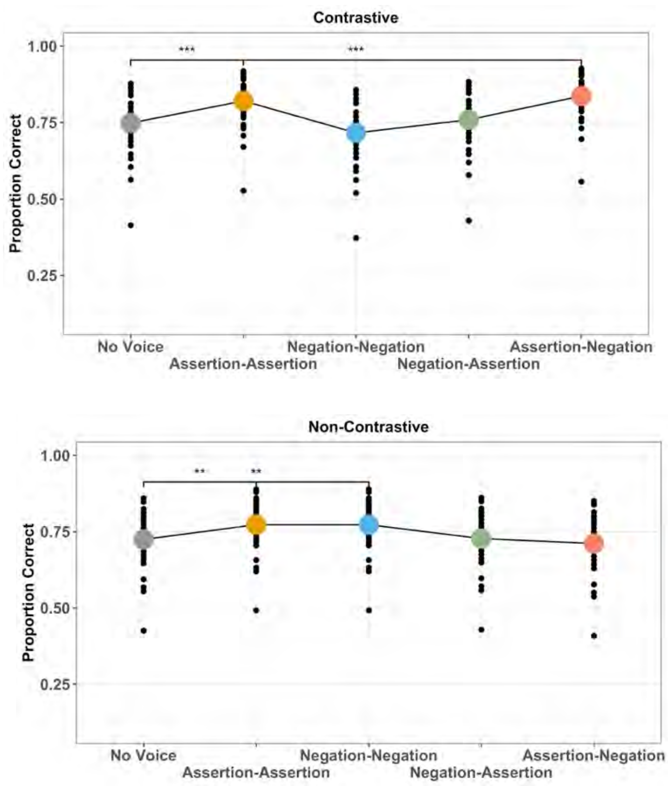


Figure 2 (b). Proportion of correct recall for C and NC path conditions. Pairwise comparison with No Voice as baseline.
(Codes: *** = p ≤ 0.001, ** = p ≤ 0.01).

References

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, 38(7), 1033–1050.

Lipton, P. (1990). Contrastive explanation. *Royal Institute of Philosophy Supplement*, 27, 247-266.

Miller, T. (2021). Contrastive explanation: A structural-model approach. *The Knowledge Engineering Review*, 36, (E14).

Mirman, D., Dixon, J.A., & Magnuson, J.S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475-494.

Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, 34, 1064–1092.

Talmy, L. (2000). Toward a cognitive semantics, vol. 1. *MIT Press*.

Supplementary materials: https://osf.io/ZFCBK/

Gaze behaviour during turn-taking in dyadic avatar-mediated conversations

Malin Spaniol[1], Alicia Janz[2], Mathis Jording[3], Martine Grice[2], Kai Vogeley[1,3]
[1] Department of Psychiatry, University Hospital Cologne, Cologne, Germany,
[2] Phonetics Laboratory, University of Cologne, Cologne, Germany
[3] Cognitive Neuroscience, Institute for Neuroscience and Medicine (INM-3), Research Center Jülich, Jülich, Germany

Human communication comprises a complex and dynamic interplay of verbal and nonverbal communication channels. The investigation of this interplay therefore represents a major methodological challenge. Technical developments in interaction platforms using virtual characters provide promising tools for these investigations. Paradigms in which participants interact with algorithmically controlled agents have already enabled the investigation of individual nonverbal communication channels with the necessary experimental control. However, it is unclear how these results relate to human-human communication. A particularly interesting case is the study of turn-taking behaviour. When yielding a turn, participants tend to direct their gaze towards the eyes of the interlocutor (e.g. Brône et al., 2017; Ho et al., 2015; Kendon, 1967), as a signal and to enable a smooth turn transition (Degutyte & Astell, 2021). At the beginning of turns, participants reportedly avert their gaze from the interlocutor (e.g. Ho et al., 2015; Kendon, 1967)(but see Beattie, 1978; Streeck, 2014 for diverging reports).

Here, we present a study with a new system for human-human interactions mediated by avatars. As a proof-of-concept, we tested the generalisability of gaze patterns during turn-taking in avatar-mediated conversations. The aim of our study was to investigate whether communication patterns found in turn-taking during face-to-face interaction can be replicated in avatar-mediated communication. We invited nine gender-matched dyads (n = 18) to the University Hospital of Cologne. Participants were seated in separate rooms in front of monitors showing an avatar of the interlocutor and could hear each other via headsets. Eye movements were recorded using Tobii 4 C eye trackers and transferred to cartoon-like avatars (see Fig. 1). Participants talked to each other for thirty minutes about instructed topics. We systematically investigated the fixation duration at each turn transition as compared to the interval during the turns (see Fig. 2). Only intervals with ≥ 70% of valid eye tracking and head-tracking data were used for the final analysis. The final dataset thus included 1067 start intervals, 1225 end intervals and 1251 control intervals.
We fitted a Bayesian mixed effects logistic regression model to speaker's gaze (binary variable; directed/ averted) as a function of turn interval type (start/ control/ end; reference level "control") using the brms package (Buerkner, 2016) in RStudio. The model included dyad, subject, conversation number, and turn as random effects. A weakly informative prior of a normal distribution with μ = 0 and σ = 0.5 was used for the fixed effect. For the random effects, the build in prior was used (Student's t-distribution ν = 3, μ = 0, σ = 2,5).

There is more gaze into the avatar's eyes towards the end of a turn (estimate β_{end} = 0.8, 95%CI = [0.33,1.26]) than turn medially (estimate β_{control} = - 0.91, 95%CI = [-3.9,2.1]). There is compelling evidence for this difference P(β_{end} > 0) = 1. There is less direct gaze after the start of a turn (estimate β_{start} = -.51, 95%CI = [-0.99, -.03]) compared to turn medially. There is compelling evidence for this difference P(β_{start} < 0) = 0.98. We conclude that the data and the model support the hypothesis that gaze patterns during turn-taking in avatar-mediated interactions are comparable to what has widely been reported in the literature for natural human-human interactions (Brône et al., 2017; Cummins, 2012; Ho et al., 2015; Kendon, 1967). Additional exploratory analyses shed light on interesting variability in individual-specific gaze behaviour during the conversation in general (e.g. range of overall directed gaze on a subject level was 16 – 90%, range of mutual gaze on a dyad level was 11 – 69%).
These findings will be investigated further. Our findings indicate that using avatar-mediated communication is a promising tool to investigate the complexity of human communication.

Keywords: conversation, turn-taking, avatars, social cognition, gaze

Figure 1. Male and female avatars. Gaze into the green region was defined as direct gaze. Gaze outside this region was defined as averted gaze.

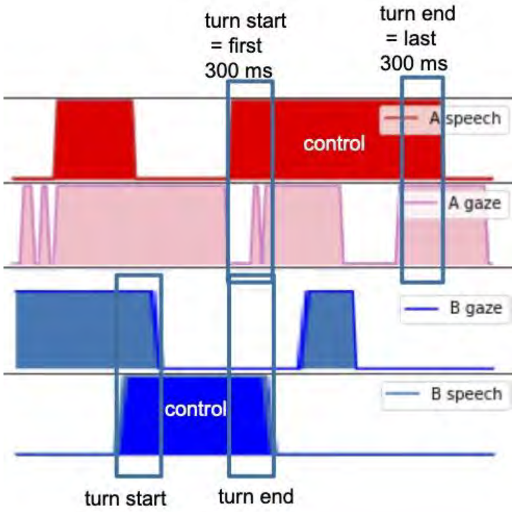


Figure 2. Segment of an exemplary turn-plot of a dyad. Turns were divided into start (300ms), control and end (300ms). Gaze behaviour was analysed in relation to interval type.

Beattie, G. W. (1978). Floor apportionment and gaze in conversational dyads. *British Journal of Social and Clinical Psychology*, *17*(1), 7–15. https://doi.org/10.1111/J.2044-8260.1978.TB00889.X

Brône, G., Oben, B., Jehoul, A., Vranjes, J., & Feyaerts, K. (2017). Eye gaze and viewpoint in multimodal interaction management. *Cognitive Linguistics*, *28*(3), 449–483. https://doi.org/10.1515/COG-2016-0119

Buerkner, P.-C. (2016). brms: An R package for Bayesian multi- level models using Stan. *Journal of Statistical Software*, *80.*(1), 1–28.

Cummins, F. (2012). Gaze and blinking in dyadic conversation: A study in coordinated behaviour among individuals. *Language and Cognitive Processes*, *27*(10), 1525–1549. https://doi.org/10.1080/01690965.2011.615220

Degutyte, Z., & Astell, A. (2021). The Role of Eye Gaze in Regulating Turn Taking in Conversations: A Systematized Review of Methods and Findings. *Frontiers in Psychology*, *12*(April), 1–22. https://doi.org/10.3389/fpsyg.2021.616471

Ho, Foulsham, T., & Kingstone, A. (2015). Speaking and Listening with the Eyes: Gaze Signaling during Dyadic Interactions. *PLOS ONE*, *10*(8), e0136905. https://doi.org/10.1371/journal.pone.0136905

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63. https://doi.org/10.1016/0001-6918(67)90005-4

Streeck, J. (2014). Mutual gaze and recognition: Revisiting Kendon's "Gaze direction in two-person conversation." *From Gesture in Conversation to Visible Action as Utterance: Essays in Honor of Adam Kendon*, 35–56. https://doi.org/10.1075/Z.188.03STR

Comparing sensory properties of words between English, Dutch, and Italian

Annika Tjuka
Max Planck Institute for Evolutionary Anthropology

The implications of language diversity for the mental representation and processing of language are of great importance to cognitive science and have received more attention in recent years (Kemmerer, 2019). Variation in word meaning is an inevitable phenomenon that needs further investigation to broaden our understanding of human minds. It is still an open question whether words in different languages that refer to the same concepts are represented similarly. Linguists have established the Concepticon database that offers standardized concept sets linked to the respective word in a given language to make judgments about historical relationships between languages (List et al., 2016). Due to expert linguists' informed decisions, this cross-linguistic database has many advantages over automatic translations of words. Cognitive scientists can benefit from a recent extension of the data – the NoRaRe database (Tjuka et al., 2021) – which includes 65 semantic properties offering information on frequency, age of acquisition, and other psycholinguistic measures for 40 languages. The present study used this cross-linguistic database to examine the question of whether the sensory properties of words are similar across English, Dutch, and Italian.

The study compared sensory modality ratings of five dimensions: haptic, visual, olfactory, gustatory, and auditory. The ratings were based on a 5-point scale and were collected for English (Lynott et al., 2020), Dutch (Speed & Brysbaert, 2021), and Italian (Vergallito et al., 2020). The word lists from these studies were mapped to the standardized concept sets in Concepticon (List et al., 2016) to enable cross-linguistic comparison. For example, the words *thunder*, *donder*, and *tuono* were mapped to the concept set 1150 THUNDER. The mapping was based on a workflow introduced for the NoRaRe database (Tjuka et al., 2021) which maps large numbers of words automatically to the concept sets. The ratings for each language pair are compared on the basis of the concept sets which occur across two word lists (Italian-English: 500 words; Italian-Dutch: 198 words; English-Dutch: 738 words). The words are mainly nouns of the basic vocabulary. The Pearson coefficients for the five sensory modalities were above R=0.7 in all language pairs as shown in Table 1. This suggests that the sensory properties of the words are perceived similarly by English, Dutch, and Italian speakers. Interestingly, subtle differences in the individual sensory modalities became apparent. For example, the correlations in the visual modality were the lowest (about R=0.7), whereas the Pearson coefficients for the auditory modality were above R=0.84 across all languages. The ratings in the gustatory modality were very similar across Dutch and English, but both differed from Italian.

The present study focused on a comparison of languages that are closely related and showed that sensory properties of words are perceived similarly across English, Dutch, and Italian speakers. However, additional data for various languages with the same rating scale need to be collected before a general claim can be made about the perception of sensory properties of words across cultures. The implications of such a large-scale study would be far-reaching because it could reveal important differences and similarities in the representation of word meanings across human minds. With the help of cross-linguistic databases, linguists and cognitive scientists can work together to answer big picture questions and generate cross-disciplinary insights about the relation between language, cognition, and culture.

Keywords: Sensory perception, Psycholinguistic properties, Language comparison, Database

Table 1: Results of the comparison of sensory modality ratings between English, Dutch and Italian based on the data in the NoRaRe database (Tjuka et al., 2021).

Language pair	**Words**	**Sensory modality**	***R***
Italian-English	500 (nouns: 380, verbs: 28, adjectives: 92)	auditory haptic visual gustatory olfactory	0.86 0.85 0.79 0.83 0.83
Italian-Dutch	198 (nouns: 139, verbs: 6, adjectives: 53)	auditory haptic visual gustatory olfactory	0.88 0.83 0.75 0.74 0.78
English-Dutch	738 (nouns: 367, verbs: 28, adjectives: 183, other: 160)	auditory haptic visual gustatory olfactory	0.84 0.77 0.73 0.9 0.83

References

Kemmerer, D. (2019). *Concepts in the brain: The view from cross-linguistic diversity*. Oxford, UK: Oxford University Press.

List, J.-M., Cysouw, M., & Forkel, R. (2016). Concepticon: A resource for the linking of concept lists. In N. Calzolari et al. (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation* (pp. 2393–2400). Portorož, Slovenia: European Language Resources Association.

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*, 1271–1291.

Speed, L. J., & Brysbaert, M. (2021). Dutch sensory modality norms. *Behavior Research Methods*, 1-13.

Tjuka, A., Forkel, R., & List, J.-M. (2021). Linking norms, ratings, and relations of words and concepts across multiple language varieties. *Behavior Research Methods*.

Vergallito, A., Petilli, M. A., & Marelli, M. (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior Research Methods*, *52*(4), 1599–1616.

Text Comprehension: Mental Representation and Assessment

Monika Tschense[1,2] & Sebastian Wallot[1,2]

[1]Leuphana University of Lüneburg
[2]Max Planck Institute for Emprical Aesthetics, Frankfurt/M.

When reading a text, a mental model is generated that serves as a base for text comprehension. As long as reading progresses, this model has to be expanded and constantly updated (e.g., Verhoeven & Perfetti, 2008). The longer a text is, the more information units come with it. However, only limited cognitive resources are available, probably most obvious for working memory capacities (cf. Carretti et al., 2009). Thus, many models of text or discourse representation postulate different processing levels when it comes to text reading and comprehension. One aim of the current study is to test whether, how and to what extent three broadly assumed building blocks of mental text representation, i.e., information units at micro level, inference level, and macro level (e.g., Kintsch & van Dijk, 1978, O'Brien & Cook, 2015), contribute to text comprehension. Furthermore, we investigated how well different comprehension items captured text comprehension after reading. Possible implications for text comprehension assessment in future studies will be discussed.

Method

400 participants were randomly assigned to read one of three short stories of comparable length, linguistic characteristics and complexity. After reading, participants were asked to summarize the text, to answer a battery of comprehension items, and to rate how interesting the story was and how much they liked reading it. For each text, 16 wh-questions and 60 yes/no statements were used to assess participants' micro- and inference-level comprehension. In order to examine text comprehension at macro level, 16 main contents were extracted for each short story, and later on used to evaluate participants' written summaries. The study was implemented as an online experiment.

Participants' answers to the comprehension items were assessed as true (1) or false (0). Furthermore, two raters evaluated the written summaries regarding the absence or presence of the 16 main contents. Subsequently, items with bad psychometric properties were excluded. First, items with accuracy rates of less than 5% or more than 95% were removed. Then joint distributions were observed by computing phi coefficients (r_ϕ) for each pair of items: While items at the the same comprehension level were supposed to correlate with each other (0.1 < mean r_ϕ < 0.9), items of different levels should be not at all or less strongly correlated (mean r_ϕ < 0.2). The remaining items were subjected to confirmatory factor analysis (CFA) and stepwise reduced until models converged. Per story, comprehension level and item type, at least three items were retained.

For each of the short stories, two different models were set up (**Figure 1**) that reflect text comprehension as (A) one-dimensional construct implemented as uni-factor model with a single comprehension factor, or as (B) multi-dimensional construct capturing all levels of text comprehension (micro level, macro level, inferences) designed as a model containing three correlated first-order factors.

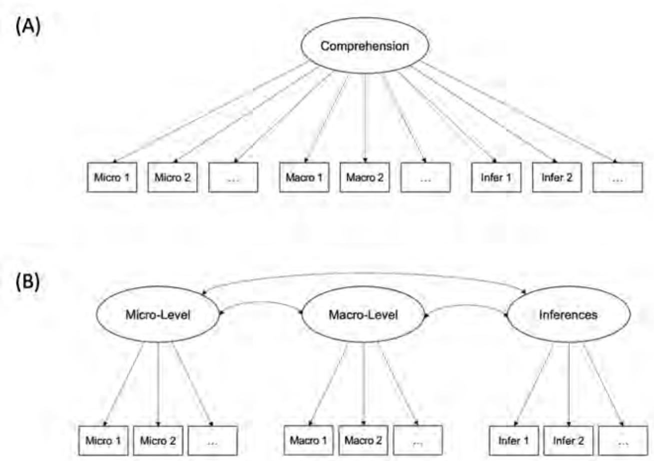
Results and Discussion

Fit indices for the the uni-factor model and the three-factor model are summarized in **Table 1**. While fit indices in general indicate that both models are a good fit for our data, the three-factor model performs even better in capturing text comprehension. This is true for all three short stories and regardless of item type. At the same time, our results showed that all three processing levels were correlated. These results suggest three related, yet distinct levels of comprehension influencing one another. Thus, all three levels should be assessed in order to adequately capture comprehension in future studies.

As the current study showed, it is of importance to adequately control the quality of items used for comprehension assessment: After reducing the initial item set by half based judgements of three independent raters, data collection was started with 16 wh-questions, 60 yes/no-statements, and 16 main contents per story. During preprocessing and data analysis, further items had to be discarded, so that a maximum of 10 wh-questions, 12 yes/no-statements and eight main contents per story could be kept. This showed that the commonly used practice of one-shot items which are largely based on experimenters' intuition for item selection rather than on theory, pre-testing or post-hoc quality control should be revisited for future studies in order to assess comprehension in a valid manner.

Keywords: text reading, text comprehension, text representation, discourse representation

Figure 1. Models for Confirmatory Factor Analysis.



Short Story	Item Type	Model	χ^2 estimate	χ^2 *df*	CFI	TLI	RMSEA
1	Statements	A: uni-factor model	150.35	119	0.90	0.89	0.05
		B: three-factor model	109.11	116	1.00	1.03	0.00
	Wh-questions	A: uni-factor model	79.55	77	0.99	0.99	0.02
		B: three-factor model	53.39	74	1.00	1.11	0.00
2	Statements	A: uni-factor model	166.73	152	0.91	0.90	0.03
		B: three-factor model	103.46	149	1.00	1.30	0.00
	Wh-questions	A: uni-factor model	116.63	90	0.78	0.74	0.05
		B: three-factor model	76.17	87	1.00	1.11	0.00
3	Statements	A: uni-factor model	223.04	170	0.78	0.75	0.05
		B: three-factor model	153.68	167	1.00	1.06	0.00
	Wh-questions	A: uni-factor model	69.89	77	1.00	1.06	0.00
		B: three-factor model	50.25	74	1.00	1.18	0.00

Table 1. Results of Confirmatory Factor Analysis. *Notes:* CFI = Comparative Fit Index, TLI = Tucker Lewis Index, RMSEA = Root Mean Square Error of Approximation.

References

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. Learning and individual differences, 19(2), 246-251.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological review, 85(5), 363.

O'Brien, E. J., & Cook, A. E. (2015). Models of Discourse Comprehension. In A. Pollatsek & R. Treiman (Eds.). Oxford Handbook of Reading (pp. 217-231). New York: Oxford University Press. 217-231.

Verhoeven, L., & Perfetti, C. (2008). Advances in text comprehension: Model, process and development. Applied Cognitive Psychology, 22, 293–301.

Structural Characteristics of Hierarchical Goal Systems from Online Field Studies

Felix Weber, fweber@uos.de
University of Osnabrück

In various subdisciplines of Cognitive Science, the "goal" construct functions to define future states of valence in humans (Vancouver & Austin, 1996) artificial agents (Stock et al., 2014), and nervous systems of biological agents (Berntson & Cacioppo, 2008; Dezfouli & Balleine, 2013) or state spaces of games or street maps. To achieve high-level goals, splitting them into increasingly more actionable subgoals and deriving action sequences from the resulting hierarchical goal systems (HGS) is an effective strategy for goal pursuit. In comparison to artificial agents, humans tend to act less structured in their goal pursuit, and there is empirical evidence for the tractability of hierarchical goal structures for humans (Bourgin et al., 2017).

To assist human learners in higher education, we have developed a web-based hierarchical goal-setting tool to help students to develop and externally maintain personal goal systems during studies.

Here, we present aggregated data from seven studies conducted from 2019 to the present, holding data from 245 participants, 555 goal systems, and 5565 goals. Based on these data, we aim to answer the following questions:

(1) Which structural characteristics, such as goal system size, branching factor, and depth, do the initial goal systems show?
(2) Which degree of variance do we find in these parameters?

On a practical level, the answers to those questions are essential for designing a software artifact, serving as technical augmentation for goal-setting, goal maintenance, and goal pursuit. On a normative level, the answers to those questions are a starting point to draw inferences on mental representations of goal systems.

Our results show an average tree size of 10.03 (SD=5.82), a minimum of 3 and a maximum of 54, an average depth of 3.27 (SD=0.78) with a minimum of 2 and a maximum of 8, and an average branching factor per node of 2.34 (SD=1.16), a minimum of 1 and a maximum of 14.

It could be argued that the representation on screens limits size, branching, and depth, but there are examples of large HGS with up to 54 goals and a depth of up to 8. These examples illustrate that the technical environment is not a limiting factor.

If according to Millers' law (Miller, 1956) the human working memory can store 7 ± 2 items, and the average HGS size in our data is 10, then the external representation in the graphical user interface on average supersedes human goal system cognition in terms of capacity. Consequently, a hierarchical goal system of the demonstrated type is suitable to augment goal-setting and meta-cognitive processes in education.

Keywords: Hierarchical Goal Systems, Digital Study Assistants, Higher Education, Goal-Setting

statistics	
Name	all
participants	245
Number of trees	555
goals	5565
min. size	3
max. size	54
av. size	10.03
SD size	5.82
min. depth	2
max. depth	8
av. depth	3.27
SD depth	0.78
min. branching	1
max. branching	14
av. branching	2.34
SD branching	1.16
leaf goals	3420
intermediate goals	1590
root goals	555
nodes	2145
branches	3420

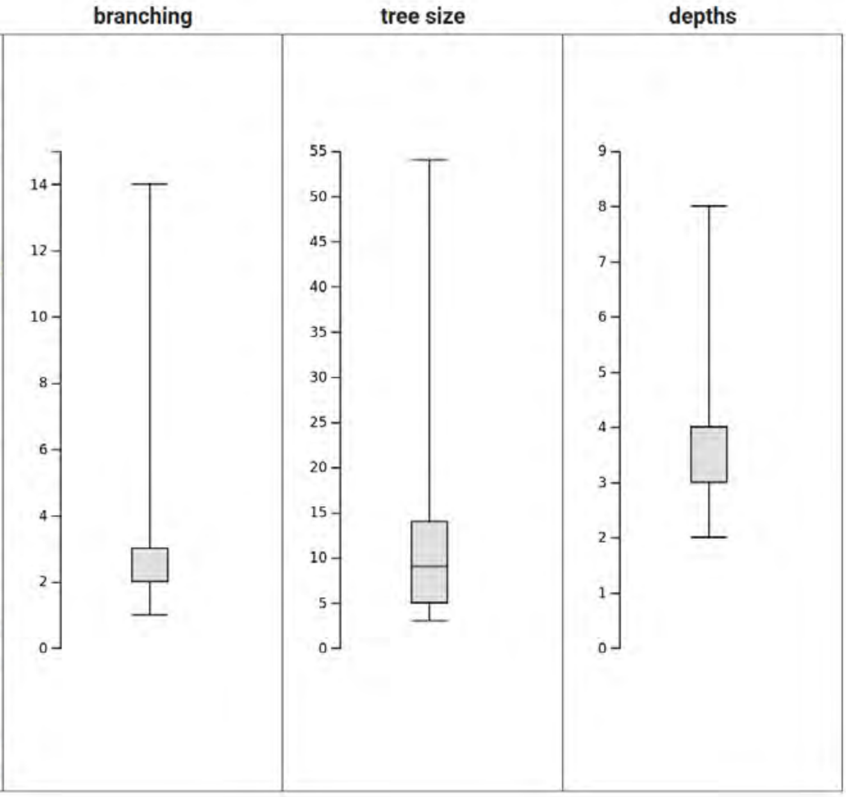


Figure 1: Structural characteristics of hierarchical goal systems from seven online studies.

References

Berntson, G. G., & Cacioppo, J. T. (2008). The neuroevolution of motivation. *Handbook of motivation science.*, 188–200.

Bourgin, D. D., Lieder, F., Reichman, D., Talmon, N., & Griffiths, T. L. (2017). The Structure of Goal Systems Predicts Human Performance. *Proceedigns of the 39th Annual Meeting of the Cognitive Science Society*, 1660–1665.

Dezfouli, A., & Balleine, B. W. (2013). Actions, Action Sequences and Habits: Evidence That Goal-Directed and Habitual Action Control Are Hierarchically Organized. *PLoS Computational Biology*, *9*(12). doi: 10.1371/journal.pcbi.1003364

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2). doi: 10.1037/h0043158

Stock, S., Günther, M., & Hertzberg, J. (2014). Generating and executing hierarchical mobile manipulation plans. *Proceedings for the Joint Conference of ISR 2014 - 45th International Symposium on Robotics and Robotik 2014 - 8th German Conference on Robotics, ISR/ROBOTIK 2014*, 605–610.

Vancouver, J. B., & Austin, J. T. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*(3), 338–375.

Meet and Greet With Knowledge: A Paradigm and Study Investigating Text Comprehension Processes With Knowledge Models Generated by T-MITOCAR

Tim Wilde, Anna Willisch, Wibke Maria Hachmann, Matthias Zaft, and Pablo Pirnay-Dummer

Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

Abstract

When taking into account personalized skill development in higher education, the encounter with knowledge - be it ones own or new knowledge - has become a key component (Heßdörfer, Hachmann, & Zaft, 2021; Pirnay-Dummer, 2020; Pirnay-Dummer, Haehnlein, Hachmann, & Umber, 2019). We derived a paradigm to investigate the gradual, cognitive processes characterised by increasing depth of text elaboration, culminating in the possible recursive reflection of text and knowledge model comprehension (possibly model revision, see Clark et al., 2012; Gentner & Stevens, 2014) at the level of propositions (associative connections between concepts). Here, this paradigm is tested to investigate the process of incremental text comprehension (Ballod, 2007; Mahn & Meyer, 2020) using re-representations of knowledge models (networks) and stepwise concept tasks to encounter in addition to text.

To provide knowledge artefacts from text that have a basis in human knowledge processing, we use the knowledge modelling software T-MITOCAR (Pirnay-Dummer, 2015; Pirnay-Dummer, Ifenthaler, & Spector, 2010). This computational linguistic heuristic uses the written language's syntax to track the associations of concepts from a text according to mental model theory (Seel, Ifenthaler, & Pirnay-Dummer, 2009; Strasser, 2010) and outputs a mathematical graph which can be visualised as a network (Pirnay-Dummer, 2010). Figure 1 shows an example network from the text used in this study.

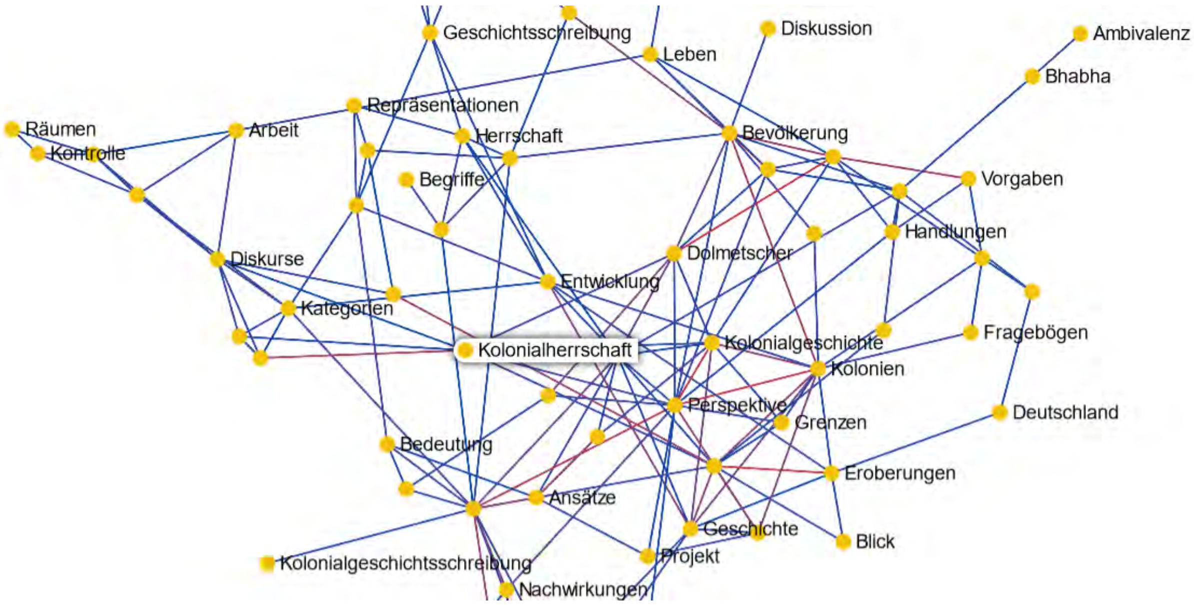


Figure 1: example of the knowledge network, interactive visualization in Sigma

This first factor is used to investigate how encountering the knowledge network alongside text fosters text comprehension. However, the presentation of knowledge networks alone might not suffice to trigger working with concepts and propositions more in depth: Based on a Grounded Theory approach to interviews with experts of diverse disciplines, who encountered their personal expertise text and knowledge model, we deduced hypotheses about the exploratory and incremental comprehension process in knowledge models, resulting in stepwise concept tasks (second factor). Here, we test these hypotheses in a 2x2 between-subject control-group design, where the use of knowledge networks (yes/no) is crossed with stepwise concept tasks (yes/no). Treatment outcome is measured as multiple choice comprehension task performance (see also Dori, Avargil, Kohen, & Saar, 2018).

Preliminary results (N=56) show higher text comprehension performance in the network and control groups, $\chi2(3) = 19.09$, $p < .001$. However, total time spent with the text content related positively to text comprehension across groups, $\chi2(3) = 13.026$, $p < .001$. In particular, the time spent reading the text for the first time related to higher comprehension performance in the groups with stepwise concept tasks than in other groups, especially for the more difficult comprehension questions, $\chi2(3) = 8.486$, $p < .05$. These results will be complemented by a larger dataset (stopping at N=80 for final analysis) and discussed with respect to mental model theory and digitally supported mentoring in more detail at the venue.

References

Ballod, M. (2007). *Informationsökonomie – Informationsdidaktik. Strategien zur gesellschaftlichen, organisationalen und individuellen Informationsbewältigung und Wissensvermittlung*. (Habilitation). Bielefeld: WBV.

Clark, M., Kim, Y., Kruschwitz, U., Song, D., Albakour, D., Dignum, S., . . . De Roeck, A. (2012). Automatically structuring domain knowledge from text: An overview of current research. *Information Processing & Management, 48(3),* 552-568.

Dori, Y. J., Avargil, S., Kohen, Z., & Saar, L. (2018). Context-based learning and metacognitive prompts for enhancing scientific text comprehension. *International Journal of Science Education, 40(10),* 1198-1220. doi:10.1080/09500693.2018.1470351

Gentner, D., & Stevens, A. L. (2014). *Mental Models*. Psychology Press.

Heßdörfer, F., Hachmann, W., & Zaft, M. (2021). Graphenbasierte Textanalyse in Lernkontexten: Technische Voraussetzungen, prototypische Szenarien, didaktische Reflexion. In H.-W. Wollersheim, M. Karapanos, & N. Pengel (Eds.), *Bildung in der digitalen Transformation (Vol. Medien in der Wissenschaft, pp. 245-256).* Münster; New York: Waxmann.

Mahn, H., & Meyer, R. J. (2020). *Vygotsky And Reclaiming Meaning*. Reclaiming Literacies as Meaning Making: Manifestations of Values, Identities, Relationships, and Knowledge, 259.

Pirnay-Dummer, P. (2010). Complete structure comparison. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 235-258): Springer.

Pirnay-Dummer, P. (2015). Linguistic analysis tools. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 427-442). New York: Guildford Publications.

Pirnay-Dummer, P. (2020). Knowledge and Structure to Teach: A Model-Based Computer-Linguistic Approach to Track, Visualize, Compare and Cluster Knowledge and Knowledge Integration in Pre-Service Teachers. In *International Perspectives on Knowledge Integration* (pp. 133-154): Brill.

Pirnay-Dummer, P., Haehnlein, I., Hachmann, W., & Umber, D. (2019). Visualizing state and change in human thought, knowledge, and attitude from written texts. *Paper presented at the Internatioal SUNBELT network conference,* Montreal, Canada.

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development, 58(1)*, 3-18. doi:10.1007/s11423-009-9119-8

Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2009). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, W. Hung, D. H. Jonassen, & J. Strobel (Eds.), *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains* (pp. 17-40). Rotterdam: Sense Publishers.

Strasser, A. (2010). A functional view toward mental representations. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds*.), Computer-Based Diagnostics and Systematic Analysis of Knowledge* (pp. 15-26). Boston, MA: Springer.

Logic and Psychology—Minding the Gap with Jean Piaget

Author: M. A. Winstanley Unaffiliated

Since Gottlob Frege's disambiguation of laws of thought, logicians and psychologists alike have adhered to a strict division of labour. This has created a gap between reasoning as a psychological phenomenon and logic as a formal discipline, a gap that has widened ever since due to the accusation of psychologism (Kusch, 2003) and evidence of reasoning experiments (Johnson-Laird, 2006, p. 17; Wason, 1966; Wason & Johnson-Laird, 1972). However, logic is a benchmark of reasoning according to the standard picture of rationality (Stein, 1996), and it is hard to accept that logic has no special role to play in reasoning (Harman, 1984). Intuitively, logic and reasoning are connected, and attempts have been made to bridge the gap in mind (e.g., Haack, 1978; Hanna, 2006; Jacquette, 2003; Leech, 2015; e.g., Stenning & van Lambalgen, 2008; van Benthem et al., 2007). However, the division of labour is often eroded, making such attempts vulnerable to the accusation of psychologism.

Jean Piaget (Piaget & Beth, 1966, sec. 42 II) developed a compass to navigate the hazards of psychologism. Metaphorically, he made logic the mirror of thought rather than vice versa (Piaget, 2001, p. 27); however, a metaphor is poor a substitute for explanation. Logic traditionally understood represented an anomaly to the structuralist tendencies Piaget discerned in neighbouring scientific disciplines, and, intent on reconciliation, he maintained that an interpropositional grouping is the operational structure that lies '"beneath"' (Piaget, 1970, p. 31) propositional logic, so that 'logic is the axiomatic of the operational structures whose real functioning psychology and sociology of thought study.' (Piaget & Grize, 1972, p. 15 my translation) Piaget alluded to properties of the interpropositional grouping implicit in various axiomatisations of propositional logic (Piaget & Grize, 1972, secs 34 & 35); however, he would have had to demonstrate concretely how propositional logic rises to the surface from this operational structure beneath to fully realise his intention.

Contemporary logicians have more liberal views than their predecessors on what counts as evidence for logical theories (Hjortland, 2019; Martin & Hjortland, 2021), and I have argued that psychology is one possible source (Winstanley, 2021). In this paper, I will indicate how propositional logic can arise from the interpropositional grouping.

Propositional calculus induces a Boolean algebra on equivalence classes of sentences. Via the relation of logical equivalence, the sentences of propositional calculus separate into equivalences classes, and the operations ∨, ∧, ~ on these sentences can be interpreted in terms of ∪, ∩, ' on these classes. Consequently, it makes unequivocal sense to denote classes of sentences by propositions. The null and all classes, denoted 0 and 1 respectively, are elements of Boolean algebras; in fact, they are the identity elements of ∪, ∩, respectively. Moreover, p∧~p=0 and p∨~p=1. From a logical viewpoint, 0 and 1 thus represent classes of contradictions and tautologies. Material implication, abbreviated p→q, is defined ~p∨q. Now, 1=~(p∨q)∨(p∨q)=~(q∨p)∨(p∨q)=(q∨p)→(p∨q) and 1=~q∨q=(~q∧~q)∨q=~(q∨q)∨q=(q∨q)→q.

Since (q∨p)→(p∨q) and (q∨q)→q are tautologies, they are formally true. Moreover, they feature in popular axiomatisations of propositional logic. According to (Halmos, 2016, p. 22), many properties of propositional calculus have algebraic alter egos. Constructing axiom schemata for propositional logic thus appears to consist in discovering a minimal subset of tautologies, from which 1, the class of tautologies, can be derived.

The interpropositional grouping is an operational structure, and logic is inherent in mathematical structures (Shapiro, 2014). Moreover, Piaget based the interpropositional grouping on Boolean algebra (Grize, 2013; Piaget, 1970, n. 9 chapter II). In this paper, I argue that axiom schemata of propositional logic not only have a Boolean alter ego but arise from the interpropositional grouping, the operational structure of thought, beneath.

Keywords: Propositional Logic, Propositional Reasoning, Psychologism, Interpropositional Grouping, Structuralism

References

Grize, J.-B. (2013). Operatory Logic. In B. Inhelder, D. de Caprona, & A. Cornu-Wells (Eds.), *Piaget Today* (electronic resource, pp. 149–164). Taylor and Francis.
Haack, S. (1978). *Philosophy of Logics*. Cambridge University Press.
Halmos, P. R. (2016). *Algebraic Logic* (Dover Edition). Dover Publications Inc.
Hanna, R. (2006). *Rationality and Logic*. MIT Press.
Harman, G. (1984). Logic and Reasoning. In H. Leblanc, E. Mendelson, & A. Orenstein (Eds.), *Foundations: Logic, Language, and Mathematics* (pp. 107–127). Springer Netherlands. https://doi.org/10.1007/978-94-017-1592-8_7
Hjortland, O. T. (2019). What Counts as Evidence for a Logical Theory? *The Australasian Journal of Logic*, *16*(7), 250–282. https://doi.org/10.26686/ajl.v16i7.5912
Jacquette, D. (Ed.). (2003). *Philosophy, Psychology, and Psychologism: Critical and Historical Readings on the Psychological Turn in Philosophy* (UniM INTERNET resource; Vol. 91). Kluwer Academic Publishers.
Johnson-Laird, P. N. (2006). *How We Reason*. Oxford University Press.
Kusch, M. (2003). Psychologism and Sociologism in Early Twentieth-Century German-Speaking Philosophy. In *Philosophy, Psychology, and Psychologism* (pp. 131–155). Springer, Dordrecht. https://doi.org/10.1007/0-306-48134-0_6
Leech, J. (2015). Logic and the Laws of Thought. *Philosophers' Imprint*, *15*(12), e1-E27.
Martin, B., & Hjortland, O. T. (2021). Evidence in Logic. In M. Lasonen-Aarnio & C. Littlejohn (Eds.), *The Routledge Handbook of the Philosophy of Evidence* (Preprint). Routledge. https://www.academia.edu/42231700/Evidence_in_logic
Piaget, J. (1970). *Structuralism* (C. Maschler, Trans.). Basic Books Inc.
Piaget, J. (2001). *The Psychology of Intelligence* (M. Piercy & D. E. Berlyne, Trans.). Routledge.
Piaget, J., & Beth, E. W. (1966). *Mathematical Epistemology and Psychology* (W. Mays, Trans.; Softcover reprint of hardcover 1st edition, Vol. 12). Springer Netherlands. DOI 10.1007/978-94-017-2193-6
Piaget, J., & Grize, J.-B. (1972). *Essai de logique opératoire* (2e éd. du Traité de logique, essai de logistique opératoire (1949), Vol. 15). Dunod.
Shapiro, S. (2014). *Varieties of logic*. OUP Oxford.
Stein, E. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Clarendon Press.
Stenning, K., & van Lambalgen, M. (2008). *Human Reasoning and Cognitive Science*. MIT Press.
van Benthem, J., Hodges, H., & Hodges, W. (Eds.). (2007). Logic and Cognition. *Topoi*, *26*(1).
Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New Horizons in Psychology* (pp. 135–151). Penguin Books.
Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content* (Vol. 86). Harvard University Press.
Winstanley, M. A. (2021). A Psychological Theory of Reasoning as Logical Evidence: A Piagetian Perspective. *Synthese*. https://doi.org/10.1007/s11229-021-03237-x

PLENARY SESSIONS

Panel Discussion
Cognitive Science - Past, Present, and Future

Andrea Bender
University of Bergen, Norway

Panel members: Gregor Schöner (chairman of the GK) and the keynote speakers (Marcel Brass, Seana Coulson and Dedre Gentner)
Host: Evelyn C. Ferstl

A few years back, an article in Nature started a heated debate about whether the endeavor of a unified, multidisciplinary field of Cognitive Science had failed. Nunez et al., (2019) argued that most of CogSci was dominated by psychology and true interdisciplinarity had not been achieved at all. At the same time, the amazing advances in AI led to more public awareness of the field, and Cognitive Science has become trendy.

The Nunez article is written from a US perspective, of course. In the German Society of Cognitive Science we also face a number of challenges: many researchers prefer publishing in their "own" disciplinary journals, the identification with CogSci is going down, the university system does not readily provide career paths for interdisciplinary scholars, and the advance of AI and big data puts attempts to understand cognition, both human and artificial, in the background. Despite these challenges, the few Master programs for CogSci in Germany are extremely popular and the field becomes more and more recognized in the public sphere.

To discuss these and related issues, Andrea Bender (Bergen, Norway) will start off the panel discussion about the future of Cognitive Science with a short talk on her views as the editor of the journal *Topics in Cognitive Science*. Further panel members are Gregor Schöner, the chairman of the GK, and the keynote speakers, Dedre Gentner, Seana Coulson and Marcel Brass.

Presidential Adress & Award Ceremony

How Can We Explain Cognitive Phenomena?

Albert Newen
Ruhr-University Bochum

Cognitive science is a multidisciplinary endeavour which aims to understand cognitive phenomena. Despite the official commitment to the relevance of multiple disciplines there is still a tendency to favour a unitary type of explanation, e.g. it started with symbolic information processing, then parallel distributed processing was favoured before neural networks entered the game. This was coming with a shift to a dominance of neuroscientific explanations. After the decade of the brain, we are now entering the phase in which computational modelling based on deep learning (and other AI strategies) or the framework of predictive processing seems to allow us to model all facets of the mind. I argue that cognitive science with its original commitment to multidisciplinary is still the right strategy to understand the mind. This will be illustrated by some examples.

Introducing the New Appearance of the Society

Barbara Kaup
University of Tübingen

In my presidential address I will introduce the new logo, the new flyer and the new website of the society. The website will be launched shortly after the conference. Flyers will be distributed to the audience during the presidential address.

Best Paper Award Presentation
To be announced

Author Index