13th Biannual Conference of the German Cognitive Science Society

kogwis2016.spatial-cognition.de

Proceedings

Thomas Barkowsky
Zoe Falomir Llansola
Holger Schultheis
Jasper van de Ven (Eds.)

Conference Organization

Thomas Barkowsky (chair)

Zoe Falomir Llansola, Holger Schultheis, Jasper van de Ven (local organizing team)

Gracia Kranz (administration)

Alexander Mittelstorb (technical support)

Program Committee

Anna Belardinelli, University of Tübingen

Sven Bertel, Bauhaus-Universität Weimar

Simon J. Büchner, University College Freiburg

Michele Burigo, Bielefeld University

Martin Butz, University of Tübingen

Claus Christian Carbon, University of Bamberg

Lewis Chuang, MPI for Biological Cybernetics

Frank Dylla, University of Bremen

Evelyn Ferstl, Albert-Ludwigs-Universität Freiburg

Paolo Fogliaroni, Vienna University of Technology

Christian Freksa, University of Bremen

Klaus Gramann, TU Berlin

Johannes Haack, Universität Potsdam

Marc Halbruegge, TU Berlin

Kai Hamburger, Justus-Liebig-University Giessen

Elisabeth Hein, University of Tübingen

Georg Jahn, TU Chemnitz

Martin Christof Kindsmüller, Universität Hamburg

Markus Knauff, Justus-Liebig-Universität Giessen

Stefan Kopp, Bielefeld University

Kai-Uwe Kuehnberger, University of Osnabrück

Hanspeter Mallot, University of Tübingen

Mark May, Helmut-Schmidt University

Tobias Meilinger, MPI for Biological Cybernetics

Franz-Benjamin Mocnik, Vienna University of Technology

Ulrike Susanne Pompe, University of Stuttgart

Felix Putze, University of Bremen

Marco Ragni, University of Freiburg

Kai-Florian Richter, University of Zurich

Bettina Rolke, University of Tübingen

Jascha Rüsseler, University of Bamberg

Nele Russwinkel, Technische Universität Berlin

Ute Schmid, University of Bamberg

Frieder Stolzenburg, Hochschule Harz

Anna Strasser, Humboldt-Universität zu Berlin

Thora Tenbrink, Bangor University

Manfred Thüring, Technische Universität Berlin

Sabine Timpf, University of Augsburg

Constanze C. Vorwerg, University of Bielefeld

Diedrich Wolter, University of Bamberg

Contents

Keynote Lectures 1

1 Architectural Design, Complexity, and Cognition
Ruth Conroy Dalton

3 Imagination and Mental Imagery
Jim Davies

5 Mental Autonomy and Mental Action
Thomas Metzinger

7 Acquiring spatial knowledge from different sources and perspectives: abilities, strategies, representations
Francesca Pazzaglia

9 Semantic typology and the Sapir-Whorf hypothesis in computational perspective
Terry Regier

Oral Presentations 11

O1: Language 11

11 The Role of the Center-of-Mass in Evaluating Spatial Language
Thomas Kluth, Michele Burigo, Holger Schultheis, and Pia Knoeferle

15 'Fragen Sie Ihren Arzt oder Apotheker!' How grammatical gender influences representations of discourse referents
Evelyn C. Ferstl, Lena Dietsche

19 Time pressure effects on semantic speech-gesture coordination
Kirsten Bergmann, Stefan Kopp

O2: Place recognition and Localization 23

23 Wall Distance as a Cue in Human Place Recognition
Marc Halfmann, Viktoria Prozmann, Nina Walker, and Hanspeter A. Mallot

27 A Maximum-Likelihood Approach to Place Recognition from Distant Landmarks
Hanspeter A. Mallot, Stephan Lancier

31 Self-Localization Accuracy and Spatial Ambiguity of Humans and Robots in a Complex Building
Rul von Stülpnagel, Vincent Langenfeld, Christoph Hölscher

O3: Categorization and Spatial Representation 35

35 Modelling different strategies in mental rotation
Alexander Lotz, Nele Russwinkel

39 Rule Learning from Incremental Presentation of Training Examples: Reanalysis of a Categorization Experiment
Christina Zeller, Ute Schmid

43 Multisensory Conflict yields Adaptation in Peripersonal and Extrapersonal Space
Johannes Lohmann, Martin V. Butz

47 Matching Matchboxes: Co-actors Create Non-conventional Communication Systems for Joint Action
Cordula Vesper, Tiffany Morisseau, Günther Knoblich, Dan Sperber

O4: Embodied Cognition 51

51 Rethinking the Keystroke-Level Model from an Embodied Cognition Perspective
Marc Halbrügge

55 Communicative signaling and self-other distinction: Next steps for an embodied hierarchical model of dynamic social behavior and cognition
Sebastian Kahl, Stefan Kopp

59 Motor expertise facilitates the cognitive evaluation of body postures: An ERP study
Dirk Koester, Thomas Schack, Iris Güldenpenning

63 Optimum statistical representation obtained from an intermediate feature level of the visual hierarchy
Noshaba Cheema, Lex Fridman, Ruth Rosenholtz, Christoph Zetzsche

O5: Spatial Perception **67**

67 Auditory stimulus detection partially depends on visuospatial attentional resources
Basil Wahn, Supriya Murali, and Peter König

71 Recognition of self-performed but visually unfamiliar dance-like actions from point-light displays
Bettina E. Bläsing

77 Space Constancy across Large Saccades
Bruce Bridgeman

O6: Computational Modeling and Assistance **83**

83 Modeling Kitchen Knowledge with LTM^C
Marc Halbrügge, Holger Schultheis

87 Model-driven Interaction Strategies of a Dialog System for Navigation and Information
Felix Putze, Elias Bordolo, Tanja Schultz

O7: Predictive Processing **91**

91 What are the contents of representations in predictive processing?
Wanja Wiese

95 Complementarity Between the Global Workspace Theory and the Sensorimotor Theory
Kinga Jeczminska

Symposia 99

S1: Formal and Cognitive Reasoning **99**

99 Formal and Cognitive Reasoning Abstracts
Christoph Beierle, Gabriele Kern-Isberner, Marco Ragni, Frieder Stolzenburg, Francois Bry, Christian Freksa, Markus Knauff

S2: Spatial Representation and Processing - What Information Do We Need? **103**

103 Effects of ageing on landmark recognition
Ramona Grzeschik, Ruth Conroy Dalton, Anthea Innes, Jan Wiener

105 The influence of extra-linguistic information on spatial language
Michele Burigo

109 Adjusting our view on perspective taking: Scalable representation structures and reference frames
Holger Schultheis

111 The decomposition of navigation behavior into simple tasks
Jascha Gruebel, Tyler Thrash, Victor Schinazi, Christoph Hölscher

113 The integration of room views
Chantal Horeis, Celia Foster, Katsumi Watanabe, Heinrich H. Bülthoff, Tobias Meilinger

115 What is Orientation?
Jakub Krukar, Angela Schwering

S3: Cognition and Manual Action **119**

119 Digit Position and Force Synergies during Unconstrained Grasping
Abdeldjallil Naceri, Alessandro Moscatelli, Robert Haschke, Marco Santello, Marc Ernst

121 Anticipating Object Interaction with the Eyes and with the Hands: Perceptual and Planning Aspects
Anna Belardinelli, Stefanie Blumenschein, Martin V. Butz

123 The Neurophysiological Interaction between Working Memory and Grasping Movements
Rumeysa Gunduz-Can, Thomas Schack, Dirk Koester

127 Distinct Effects of Visuomotor Priming on Action Preparation and Motor Programming
Christian Seegelke, Iris Güldenpenning, Julian Dettling, Thomas Schack

129 Interactions of Cortical Networks for Object Recognition and Object Grasping
Marc Himmelbach, Mareike Gann

S4: Dynamics of Sketching and Sketch Understanding (DySket) **131**

131 Dynamics of Sketching and Sketch Understanding Abstract
Ahmed M. H. Abdel-Fatah, Haythem O. Ismail, Kai-Uwe Kühnberger, Malumbo Chipofya, Stefan Schneider, Oliver Kutz, Kirsten Bergmann, Zoe Falomir Llansola

S5: Social Perception **135**

135 Summary and Abstracts of the Symposium
Tobias Schlicht, Shaun Gallagher, J. Suilin Lavelle

S6: Mental Files in Cognitice Science: Core cognition, Concepts and Mindreading **139**

139 Summary of Abstracts of the Symposium
Albert Newen, Francois Recanati, Josef Perner, Joulia Smortchkova

S7: PROSOCRATES: Problem Solving, Creativity and Spatial Reasoning in Cognitive Systems **143**

143 Insight and evolution
Anna Fedor, István Zachar, András Szilágyi, Michael Öllinger, Harold P. de Vladar, Eörs Szathmáry

145 On the role of physical space for commonsense problem solving
Christian Freksa

147 Experience Understanding and Creativity
Bipin Indurkhya

151 Towards Re-representation in Cognitive Systems
Ana-Maria Olteţeanu

155 Challenges and Directions for Making Cognitive Systems Creative
Kai-Uwe Kühnberger

159 Qualitative reasoning models to help solving spatial ability tests
Zoe Falomir Llansola

Tutorials 163

163 T1: Detecting and Discouraging Non-Cooperative Behavior in Online Rating Tasks
Jana Häussler, Tom Juzek

165 T2: Kant and Cognitive Science
Tobias Schlicht

167 T3: Workshop on Creativity
Bipin Indurkhya

169 T4: Introduction to Cognitive Modeling with ACT-R
Nele Rußwinkel, Sabine Prezenski, Marc Halbrügge, Stefan Lindner

171 T5: Bayesian Data Analysis: Main Ideas, Practices, and Tool
Michael Franke, Fabian Dablander

Poster Presentations 173

173 Representational Dynamics of Problem Solving in Imagery: An Exploratory Case Study
Benjamin Angerer, Cornell Schreiber, Stefan Schneider

175 Twin compatibilities: Studying spatial cognition with social Simon stimuli
Pamela Baess, Christina Bermeitinger

181 The impact of sleep on navigation and consolidation of survey knowledge
Matthias-Philipp Baumann, Wiebke Schick, Hanspeter A. Mallot

183 Moral Decision Making in Autonomous Vehicles
Lasse T. Bergmann, Silja Timm, Max Wächter, Anke Dittmer, Felix Blind, Carmen Meixner, Larissa Schlicht, Anja Faulhaber, Juhee Jang, Aalia Nosheen, Simeon Kraev, Max Räuker, Leon Sütfeld, Achim Stephan, Peter König, Gordon Pipa

185 Decoding spatial auditory attention using ear EEG
Martin G. Bleichner, Bojana Mirkovic, Stefan Debener

187 Learned knowledge about the co-actor's behavior influences performance in a joint visuomotor task
Artur Czeszumski, Ernesto Andrés López Montecinos, Chiara Carrera, Anette Aumeistere, Ann Xavier, Basil Wahn, Peter König

189 Quick and sustained inhibition of distractor elicited response activation in task switching
Philipp Dehmel, Kerstin Jost, Aquiles Luna-Rodriguez, Mike Wendt, Thomas Jacobsen

191 The Role of Facial Mimicry in Cognitive and Emotional Empathy and Effects of Autistic Traits
Hanna Drimalla, Isabel Dziobek

193 Towards Improving Users' 3D Spatial Skills using a Qualitative 3D Descriptor and a Computer Game
Zoe Falomir Llansola, Eric Oliver

195 The Cortical Network of Usability Evaluations for Unknown Tools
Mareike Gann, Marc Himmelbach

197 Neural Correlates of Semantic Expectation in a Conversation - A Wireless EEG Study of the N400 Effect
Tatiana Goregliad Fjaellingsdal, Esther Ruigendijk, Stefan Scherbaum, Martin G. Bleichner

199 The interaction dynamics of meta-control parameters and congruency proportion in spatial set shifting
Tobias Grage, Simon Frisch, Stefan Scherbaum

201 Landmark preference during route encoding and retrieval
Karoline Greger, Rebecca Albrecht, Rul von Stülpnagel

203 The impact of a humanoid robot's action-selection strategy on humans' perceived naturalness of interaction – A User Study with NAO Playing Rock-Paper-Scissors
Tobias Jakubowitz, André Kowollik, Ute Schmid

205 Semantics of Persian Spatial Term Jelo Based on Principled Polysemy Model
Marjan Daneshvar Kashkooli

209 Challenging the distinction between presupposition holes and plugs
Pritty Patel-Groz, Gerhard Jaeger, Matthias Holweger, Nadina Kiss

211 Learning to cope with uncertainty during the acquisition of mental models
Johanna Renker, Gerhard Rinkenauer

213 Dissociating components of cognitive flexibility in semantic space: Continuous measures dynamic modeling and clinical assessment
Stefan Scherbaum, Tobias Grage, Moritz Walser, Katrin Hummel, Maja Dshemuchadse

215 Language Cues in the Formation of Hierarchical Representations of Space
Wiebke Schick, Marc Halfmann, Gregor Hardiess, Hanspeter A. Mallot

217 Co-representation of Others' Spatial Task Constraints in Joint Action
Laura Schmitz, Cordula Vesper, Natalie Sebanz, Günther Knoblic

219 I-CARE: Individual Activation of People with Dementia
Tanja Schultz, Felix Putze, Timo Schulze, Ralf Mikut, Wolfgang Doneit, Andreas Kruse, Anamaria Depner, Ingo Franz, Marc Aurel Engels, Philipp Gaerte, Dietmar Bothe, Christof Ziegler, Irene Maucher, Michael Ricken, Todor Dimitrov, Joachim Herzig, Keni Bernardin, Tobias Gehrig, Jana Lohse, Marion Adam, Monika Fischer, Massimo Volpe, Clarissa Simon

221 Let's decide together! Joint delay decision-making improves delay discounting
Diana Schwenke, Maja Dhesmuchade, Cordula Vesper, Martin G. Bleichner, Stefan Scherbaum

223 Do Pitch and Space Share Common Code?
Pulkit Singhal, Aditya Agarwala, Priyanka Srivastava

225 Recent response conflict modulates early distractor processing
Michael Sprengel, Markus Hofmann, Mike Wendt, Aquiles Luna-Rodriguez, Sascha Tamm, Thomas Jacobsen, Arthur M. Jacobs

227 Emotional Effects on Time Estimates during Intervals up to 5s
Anna Katharina Trapp, Manfred Thüring

229 Predicting patterns in navigator-driven placement of landmarks for future wayfinding with Space Syntax
Rul von Stülpnagel, Christoph Hölscher

Doctoral Symposium **231**

231 3D-Shape-Perception studied exemplarily with Tetrahedron and Icosahedron as prototypes of the polarities Sharp versus Round
Iris Sauerbrei, Jörg Troja, Erich Lehner

237 Representation of Wayfinding and Perception Abilities in Agent-Based Models
Erik Andresen

241 Inductive Learning of Categories – Between Cognitive Modeling and Machine Learning
Christina Zeller

245 Modelling Human Navigation: Cognitive Aspects of Obstacle Avoidance
Juan Purcalla Arrufi

249 Cognitive Complexity of Number Sequence Completion Problems: Evidence for Human Heuristics
Martin Hillebrand

253 Updating of Spatial Representations: Two pilot ERP Experiments
Efrosini (Frosso) Charalambous

257 The comprehension of verbal jokes: A visual-world study
Laura Israel, Evelyn C. Ferstl

261 Cognitive Sciences Strategies for Futures Studies (Foresight)
Ahmad Mahdeyan, Ebrahim SoltaniNasab

Index of Authors **265**

Architectural Design, Complexity, and Cognition

Ruth Conroy Dalton

In earlier work by Dalton (and collaborators), it was suggested that there are three distant ways in which cognition takes place in relation to architectural design: 1) the impact of architectural structure, function and form on human perception, cognition and behaviour; 2) the impact of cognitive factors on the design of architectural structures; 3) the means of interaction and communication between the architect and building-user perspectives. Architectural cognition embraces all of these three types. This talk will focus on architectural cognition and complexity and, in particular, will on the lay-person/building-user's perception of complex architectural environments. It can be argued that complexity in architectural design has two principal impacts on the building user: first in their aesthetic appraisal and second in their understanding of the building layout and subsequent wayfinding through a complex environment. This keynote will cover both of these aspects of architectural but will focus on/describe experiments on the latter.

Imagination and Mental Imagery

Jim Davies

What is imagination, and how is it different from mental imagery? In this talk, Jim Davies will discuss what science shows about how imagination works, with a focus on his computational modeling of imagination and his neural model of visual mental imagery.

Mental Autonomy and Mental Action

Thomas Metzinger

I will have two central goals in the first part of this talk, which explores the relevance of latest research on mind-wandering for theories of consciousness. First, conceptually, and in opposition to what many philosophers following Descartes and Kant traditionally have liked to believe, I will argue for the claim that conscious thought actually is a subpersonal process, only rarely a form of mental action, but rather an unintentional form of mental behavior, and demonstrably for more than two thirds of our conscious life-time. The paradigmatic, standard form of conscious thought is non-agentive, it lacks veto-control, and involves an unnoticed loss of epistemic agency and goal-directed causal self-determination on the level of mental content. Second, I present an empirical hypothesis: There will be a detectable self-representational blink (SRB), a small time window I which we are blind to ourselves, namely, when shifting from one phenomenal self-model or "unit of identification" (UI) to the next. Alluding to the well-studied phenomenon of the attentional blink (Raymond, Shapiro, and Arnell, 1992, Shapiro, Raymond, and Arnell, 1997), the notion of a "self-representational blink" refers to the fact that we are typically not able to consciously experience the actual moment of transition from mindful, present-oriented self-awareness to the identification with the "protagonist" of a daydream, the content of the self-model in autobiographical planning, etc. Phenomenologically, the SRB is characterized by a brief loss of self-awareness, followed by an involuntary shift in the phenomenal UI; functionally, we can describe it as a failure of attentional and/or cognitive self-control. The empirical prediction is that subjects should be blind to self-related stimuli during the SRB, and my main hope is that the audience can help in developing novel experimental paradigms to test this hypothesis.

If time allows, I will also take a closer look at the concept of "mental action" in the second part. Can we conceptually accommodate mental actions under a predictive processing approach? My main positive claim will be that mental action is the predictive control of effective connectivity, where what is predicted is the epistemic value of states integrated into the phenomenal self-model under counterfactual outcomes.

Metzinger, T. (2013). The Myth of Cognitive Agency: Subpersonal thinking as a cyclically recurring loss of mental autonomy. Frontiers in Psychology, 4, 931.

Metzinger; T. (2015). M-Autonomy. Journal of Consciousness Studies, 22 (11-12), 270-302. Special Issue edited by Mihretu P. Guta and Sophie Gibb: Insights into the First-Person Perspective and the Self - An Interdisciplinary Approach.

Acquiring spatial knowledge from different sources and perspectives: abilities, strategies, representations.

Francesca Pazzaglia

The ability to acquire spatial knowledge is very important in everyday life, and it has been very important to the survival of our own and other species. We acquire spatial knowledge starting from a variety of sensory inputs (e.g. vision, vestibular sense, kinesthesis, motor afference) and relative encoding processes, which lead to the construction of an internal representation of the environment on which we rely to perform various spatial tasks, such as retracing a route, estimating distances and directions, or drawing a map (Wolbers & Hegarty, 2010).

An environment can be experienced (and described) from different perspectives (route or survey perspectives; Taylor & Tversky, 1992; Pazzaglia et al., 2012) and in different ways, by moving around in it, inspecting it from above, looking at a map, or listening to a verbal description. Learning experience and perspective can influence the resulting cognitive map and, as a consequence, the performance of spatial tasks (e.g. Thorndyke & Hayes-Roth, 1982). So can a number of other individual factors, including spatial abilities, visuospatial working memory (VSWM), sense of direction (SOD), and spatial representation preferences. In my paper I first introduce concepts such as SOD and spatial strategies, and then go on to describe a number of instruments widely used to assess these variables. Then I review the main outcomes of several studies based on the use of these instruments, The goal is to shed light on how interactions among these variables affect performance in spatial tasks. I also examine the role of these factors in conjunction with that of spatial ability and VSWM in determining individual differences in the performance of way-finding tasks, map learning and spatial text processing.

Semantic typology and the Sapir-Whorf hypothesis in computational perspective

Terry Regier

Why do languages have the semantic categories they do, and what do those categories reveal about cognition? Word meanings vary widely across languages, but this variation is constrained. I will argue that this pattern reflects a range of language-specific solutions to a universal functional challenge: that of communicating precisely while using minimal cognitive resources. I will present a general computational framework that instantiates this idea, and will show how that framework accounts for cross-language variation in several semantic domains. I will then address the Sapir-Whorf hypothesis - the claim that such language-specific categories in turn shape cognition. I will argue that viewing this hypothesis through the lens of probabilistic inference has the potential to resolve two sources of controversy: the challenge this hypothesis apparently poses to the widespread assumption of a universal groundwork for cognition, and the fact that some findings supporting the hypothesis do not always replicate reliably.

The Role of the Center-of-Mass in Evaluating Spatial Language

Thomas Kluth[1], Michele Burigo[1], Holger Schultheis[2], and Pia Knoeferle[3]

[1] CITEC (Cognitive Interaction Technology Excellence Cluster), Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany, {tkluth, mburigo}@cit-ec.uni-bielefeld.de
[2] Cognitive Systems Group, Department of Computer Science, University of Bremen, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany, schulth@informatik.uni-bremen.de
[3] Department of German Language and Linguistics, Humboldt University, Unter den Linden 6, 10099 Berlin, Germany, pia.knoeferle@hu-berlin.de

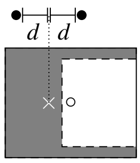
Abstract. Consider a display with a circle and a rectangle and the sentence "The circle is above rectangle." How well does the sentence describe the display? For such an acceptability rating of a spatial preposition, the location of the center-of-mass of the rectangle (reference object, RO) is assumed to play an important role. However, there is only little empirical evidence that favors the use of the center-of-mass over other possible reference points of the RO (e.g., the center-of-object). We present an empirical rating study that contrasts the center-of-mass with the center-of-object of an RO by using asymmetrical ROs. The results of the study suggest that people base their acceptability ratings on the center-of-object instead of on the center-of-mass of an RO. Computer simulations of cognitive models implementing this strategy support this view.

Keywords: spatial language; center-of-mass, center-of-object, cognitive modeling.

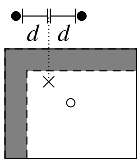
Introduction Spatial language is an important part of spatial cognition. People use spatial terms to express their mental representations of space. In this paper, we focus on the acceptability of projective spatial prepositions such as "above" or "to the left of" for describing a scene. Imagine you look at a picture that contains geometrical shapes and hear a sentence like "The circle is above the rectangle". This sentence locates the circle (*located object*, LO) relative to the rectangle (*reference object*, RO). Whether this sentence is an acceptable description of the scene depends on the relative locations of these two objects.

According to [4,5], people use two points of the RO as anchor for their acceptability ratings: The proximal point (the point of the RO that is closest to the LO) and the *center-of-mass* (CoM) of the RO. The orientations of the two imaginary lines that connect each of these two points with the LO (simplified as a single point) are called *proximal orientation* or *center-of-mass orientation*, respectively. [5] provide evidence that both of these orientations affect acceptability ratings of spatial prepositions. In this paper, we focus on the role of the CoM orientation and contrast the role of the CoM with another central point in the RO, the *center-of-object* (CoO). We define the CoO as the following point: $CoO(x,y) = \left(RO_{x0} + \frac{RO_{width}}{2}, RO_{y0} + \frac{RO_{height}}{2}\right)$, where RO_{x0} is the leftmost point of the RO and RO_{y0} is the point of the RO with the lowest y-coordinate (y-axis grows from bottom to top). The CoO coincides with the center of the bounding box of the RO (the smallest rectangle that includes all points of RO). In Figs. 1a and 1b, the bounding box is depicted as solid line, the CoO is depicted as ∘, and the CoM is depicted as ×.

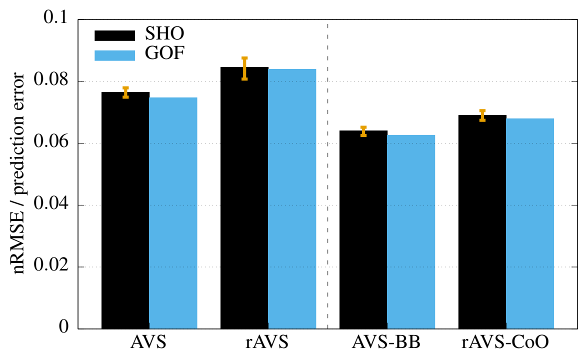
Although research on saccadic and perceptual localization has revealed that the CoM may not be the only critical point for object localization (e.g., [2,7]), the possibility of reference points other than the CoM have so far not been studied in spatial language. In most spatial language acceptability rating tasks, symmetrical ROs were used for which the location of the CoM coincides with the location of the CoO. To our knowledge, there exists only one experiment explicitly designed to dissociate the CoM from the CoO (exp. 4 from [5]). In this experiment, however, only four LOs above two ROs were tested (in total eight LOs). For ROs with a cavity on their top, the results suggest that the CoM is more important than the CoO. For ROs that have a cavity at their bottom (i.e., ROs with a flat top), the results are less clear. We conducted a study to more closely contrast the importance of the CoM with the importance of the CoO using a larger set of items (28 LOs above 4 ROs with flat tops, in total 112 LOs).



(a) Asym. “C”



(b) Asym. “L”



(c) Goodness-Of-Fit (GOF) and Simple-Hold-Out (SHO) results.

Fig. 1: (a) and (b): Stimuli used for the computational and empirical studies (dashed line = borders of the RO, solid line = bounding box of the RO, × = center-of-mass, ∘ = center-of-object, •= LO), (c): Results of model simulations.

Empirical Study We designed asymmetrical ROs for which the CoM is dissociated from the CoO (see Figs. 1a and 1b where the CoM is depicted as × and the CoO is depicted as ∘). To control for left-right biases we also included vertically mirrored versions of these ROs. We placed 28 LOs at different positions above each RO, resulting in 112 tested LOs in total (28 LOs × 4 ROs). (We also tested 4 more ROs as well as LOs placed below the ROs with the preposition “unter” (below/under). The results for these other 4 ROs and “unter” are not discussed here.)

For the predictions of the study, consider Fig. 1a. Since the two LOs (black dots) are placed at the same elevation with equal horizontal distance d to the CoM of the RO, both LOs have the same CoM orientation. Since they also have equal proximal orientations, people should rate these two LOs identically (following the reasoning by [4,5]). However, if instead the CoO is more important for the acceptability of spatial language, the right LO should be rated higher than the left LO, as it is closer to being directly above the CoO.

Each of our 34 participants saw the German sentence "Der Punkt ist über dem Objekt" ("The point is above/over the object"). After they read the sentence they had to press the space bar. Then, a picture appeared on the screen showing one RO and one LO. Participants had to rate how well the sentence described the depicted scene on a scale from 1 (sentence does not describe the picture at all) to 9 (sentence describes the picture perfectly) using the number keys above the letter keys on a keyboard. Each participant rated all LOs in a pseudo-randomized order (the same RO never appeared twice in a row.)

Results Interestingly, the CoM orientation did not have the expected effect as LOs with the same CoM orientation were rated differently: LOs that were placed above the mass of the RO were rated lower than LOs that were placed above the cavity of the RO (mean difference: 0.518, 95% confidence intervals: 0.619, 0.428). In contrast, our results suggest that people use the CoO of an RO in rating its acceptability. LOs with the same CoO orientation were rated equally: LOs on the left side of the RO received equal ratings compared to LOs on the right side of the RO (mean difference: 0.034, 95% confidence intervals: -0.101, 0.165).

We dissociated the CoM from the CoO in our stimuli, permitting us to dissociate the effects of the corresponding orientations on ratings. The results of this comparison suggested that what was previously thought to be an effect of the CoM is in fact an effect of the CoO, at least for ROs with flat tops. [5] also used asymmetrical ROs in their experiment 4 to contrast the CoO with the CoM. While they found the CoM to be more important for the RO without a flat top, they could not find such effect for the RO with a flat top.

Model Simulations Two cognitive models that compute spatial language acceptability ratings rely on the CoM for their computations: the *Attentional Vector Sum* model (AVS, [5]) and the *reversed AVS* (rAVS) model, a recently proposed modification of the AVS Model (see [1] for a motivation of the rAVS model and details of both models). According to our empirical findings, however, people seem to use the CoO instead. This is why we next present refined versions of both models.

The AVS model relies on the CoM because it computes a vector sum using all points of the RO. We modified the AVS model so that it uses all points that are in the bounding box of the RO and call this modification AVS-BB model. The bounding box is the smallest rectangle that includes all points of the RO (see solid line in Figs. 1a and 1b). In particular, the bounding box also includes the points inside the cavity of asymmetrical ROs. Computing a vector sum with all points inside the bounding box then means that the AVS-BB model relies on the CoO instead of on the CoM.[4]

The rAVS model explicitly uses the location of the CoM of the RO and can be easily modified to use the CoO instead by replacing every occurence of the CoM with the CoO. This yields the rAVS-CoO model. The AVS-BB model and the rAVS-CoO model both show the same output pattern as the empirical data: LOs above the mass are rated lower than LOs above the cavity of the RO because ratings peak above the CoO.

[4] One could also add a parameter γ that gives different attentional weights to points that are in the bounding box but outside the RO (i.e., in the cavity of the RO) compared to points that are in the RO. This would create a model that could behave like the AVS model ($\gamma = 0$) or like the AVS-BB model ($\gamma = 1$). Since this adds possibly unneeded flexibility to the model, we decided against this implementation.

We fitted all four models to the 112 empirical mean ratings by searching for values of the free model parameters that provide minimal normalized Root Mean Square Errors (nRMSE). The resulting Goodness-Of-Fit (GOF) values are plotted in Fig. 1c. An nRMSE of 0.0 means that the model is able to reproduce the empirical data exactly, while an nRMSE of 1.0 means that model output and empirical data are maximally different. As can be seen in Fig. 1c, all models can closely fit the data (GOF < 0.084). The versions of the models that use the CoO (AVS-BB and rAVS-CoO) fit the data better than the original versions (GOF < 0.068).

Since a good fit to data is a necessary but not sufficient property of a model (e.g., the good fit might be the result of overfitting due to an overly flexible model, see [3]), we also assessed the model with the simple hold-out (SHO) method proposed by [6]. [6] showed that this model selection method provides results comparable to other model selection methods. The SHO method is a cross-validation method: The data is randomly split into a training and a test set and the parameters of the model are estimated on the training set. Using the best parameters for the training set, a prediction error on the test set is computed (again an nRMSE). This is done several times with different random splits of the data. Fig. 1c shows the median prediction error of 101 SHO iterations together with their bootstrapped 95% confidence intervals. These results now clearly favor the modified versions of the models (lower SHO without overlapping confidence intervals for AVS-BB and rAVS-CoO compared to AVS and rAVS). Also, both versions of the AVS model outperform the corresponding versions of the rAVS model: AVS performs better than rAVS and AVS-BB performs better than rAVS-CoO.

Conclusion We presented an acceptability rating study of spatial prepositions using asymmetrical ROs that allowed us to explicitly contrast the importance of the CoM with the CoO. In contrast to previous literature claiming that people use the CoM of the RO as reference point ([4,5]), our results suggest that people rather select the CoO of the RO as reference point. Furthermore, we modified two cognitive models in order to implement this strategy. These modified models performed considerably better than the original models that rely on the CoM of the RO. This corroborates the importance of the CoO over the CoM.

References

1. Kluth, T., Burigo, M., Knoeferle, P.: Shifts of attention during spatial language comprehension: A computational investigation. In: Proc. of ICAART 2016, Vol. 2. pp. 213–222 (2016)
2. Nuthmann, A., Henderson, J.M.: Object-based attentional selection in scene viewing. Journal of vision 10(8), 20–20 (2010)
3. Pitt, M.A., Myung, I.J.: When a good fit can be bad. Trends Cogn Sci 6(10), 421–425 (2002)
4. Regier, T.: The human semantic potential: Spatial language and constrained connectionism. MIT Press, Cambridge, Mass. (1996)
5. Regier, T., Carlson, L.A.: Grounding spatial language in perception: An empirical and computational investigation. Journal of Experimental Psychology: General 130(2), 273–298 (2001)
6. Schultheis, H., Singhaniya, A., Chaplot, D.S.: Comparing model comparison methods. In: Proc. of the Ann Conf of the Cog Sci Soc. pp. 1294 – 1299. Cog Sci Soc, Austin, TX (2013)
7. Vishwanath, D., Kowler, E.: Localization of shapes: Eye movements and perception compared. Vision Research 43(15), 1637–1653 (2003)

"Fragen Sie Ihren Arzt oder Apotheker!" How grammatical gender influences representations of discourse referents

Evelyn C. Ferstl & Lena Dietsche

IIG / Cognitive Science
Universität Freiburg

1 Introduction

"Fragen Sie Ihren Arzt oder Apotheker!" This slogan is known to everyone who has ever watched a German TV commercial for medication. It illustrates the use of the generic masculine form in German, a language with grammatical gender. The masculine form is used whenever the gender of the person referred to is unknown or irrelevant. Although 80% of all pharmacists in Germany are female, the sentence is grammatically correct and considered standard.

Feminist linguistics has long argued that this use of person referents leads to an overestimation of male participants in unspecified groups, and that it is necessary to implement gender fair language by using alternative expressions (e.g., Ärzte (*m*) oder Ärztinnen (*f*), ÄrztInnen – *Binnen-I*) (e.g., [1]). In contrast, proponents of the traditional use argue that misunderstandings only arise when generic forms are interpreted as specific, i.e., as referring to male people.

1.1 Masculine role names: Are they generic or specific?

A number of studies have been conducted to investigate which of these two views better represents actual language use (e.g., [2-4]). Most experimental results show that grammatical gender influences the represention of the discourse information. Estimates of the percentages of people present, for instance, at a fitness studio fire described in a newspaper article, increase with gender fair language ([2]). Grammaticality judgments for sentences in which women are anaphorically related to a previous male, generically intended discourse referent reflect the power of the grammatical information ([3]), and reading times increase ([4]).

These and many other studies show that generic use leads to overestimation of male participation and that it induces processing difficulties. However, they do not yet confirm that the content of the resulting representation indeed includes the gender of the protagonists. Thus, an open question is whether hearing or reading sentences about an "Arzt" directly induces a mental image of a male doctor.

1.2 Sentence-picture matching task

Text comprehension research provides a framework to investigate this issue. One method to directly assess the representations induced by linguistic information, i.e., the situation model of the text ([5]), is sentence picture matching. In support of an embodied cognition approach, Zwaan and colleagues ([6]) used this method to show that features are included in the situation model which are not explicitly mentioned, but which are relevant for the szenario described. Thus, after hearing the sentence "The ranger saw the eagle in the sky" it takes longer to match the picture of an eagle sitting on the perch to the sentence than the picture of an eagle with spread wings.

In this study, we applied a sentence picture matching task to investigate the representations induced by generic, masculine person referents in German. The hypothesis was that it would be more difficult to match a picture to a sentence if the natural gender of the person shown on the picture was different from the grammatical gender of the person referent in the sentence. If the male form was interpreted generically, however, a female picture after a male person referent was expected to be acceptable.

2 Methods

2.1 Design and Materials

The design of the experiment varied *Word Gender* (male, female), and *Picture Gender* (male, female) within subjects. To create 12 trials per condition, 48 German nouns were selected from Misersky et al. [7], who had collected gender bias norms for a large number of occupation names. The nouns covered a wide range of gender stereotype values (from strongly masculine to strongly feminine). To each occupation, two pictures were selected: one showed a woman of this occupation, the other a man. In addition, to each of the occupations a sentence was written in which the word was explicitly mentioned. Two versions were created: one with the female noun ("Ärztin"), one with the male noun ("Arzt"). The average sentence length was about 25 words. Example stimuli are shown in Figure 1.

In addition, 72 filler sentences were written and paired with pictures to distract from the purpose of the experiment and to balance yes- and no-responses. Some pictures showed objects, others showed semantically close same-gender distractors. Four counterbalancing lists with 120 trials per list were created.

2.2 Procedure

The experiment was implemented on the on-line web platform SoSci-Survey (www.soscisurvey.de) that guided sentence and picture presentation and collected responses and reaction times for each trial. The participants were recruited via social networks and personal invitation by the experimenter (LD). 41 participants completed the entire experiment and were included in the analyses (20 women and 21 men, 27 years of age on average). After instructions, a demographic questionnaire, and practice trials, the experimental trials were presented in individually randomized order.

The participants pressed “yes” when they thought the object or person shown in the picture was mentioned in the sentences, and “no” if they did not recognize the word. Missing responses were recorded after a cutoff of 4 seconds. After about half of the trials, participants could take a break. The average duration was 25 minutes.

Fig. 1. Sample stimuli. The sentence “Nach einem Blick auf die große Uhr am Hauptgebäude der Kaserne hob **der General** (*m*) [**die Generalin** (*f*)] den Arm und die Kompanie setzte sich in Bewegung.“ [approximate translation: “After looking at the large clock at the barracks building **the general** raised the arm and the parade got on its way.”] was followed by one of the pictures shown. The condition in which the male noun (General) was followed by the picture showing a woman general is the test case for the use of the generic masculine form.

3 Results

Figure 1 shows the descriptive data. Pictures with matching gender were accepted in a majority of the trials. When the role noun was masculine, only 41% of the trials with female pictures were accepted, indicating that the generic interpretation was not always used. Interestingly, 29% of the trials in which a male noun was followed by a female picture were also accepted.

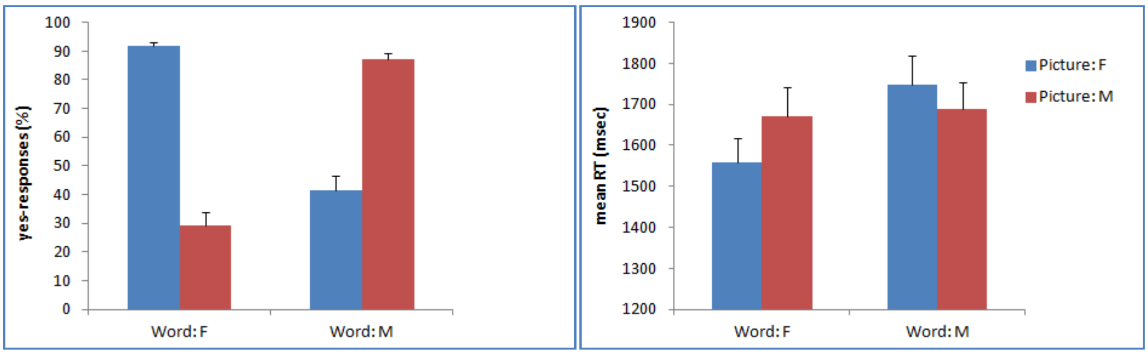


Fig. 2. Responses and reaction times (*M* and *se*) for each of the four conditions.

In an ANOVA (including age and gender) *Word* (F(1,38)=5.7, p=.02), *Picture* (F(1,38)=19.3, p<.0001), and their interaction (F(1,38)=122.1; p<.0001) were significant. Post-hoc comparisons confirmed the difference between the two inconsistent conditions (MF vs. FM; t(40)=4.3; p <.0001). There was no correlation between the proportion of yes-responses and the overall reaction times. To evaluate condition effects, reaction times for all answers (yes and no) were analysed. The longest reaction times were observed for the generic masculine condition (MF), yielding a a main effect of *Word* (F(1,38)=12.2; p=.001), and an interaction (F(1,38)=5.2, p=.03). Pairwise comparison showed an effect only for MF vs. FF (t(40)=4.6; p<.0001).

4 Conclusions

The experiment shows that a generic interpretation masculine nouns is not automatic. Picture verification responses and reaction times confirmed that it is difficult to interpret women to be referred to by a masculine noun. More fine-grained analyses, taking into account response categories and gender stereotype bias, will shed light on the interactions between grammatical form and general world knowledge.

The results are clear-cut, even though the experiment was conducted on a web platform which does not provide much control, both in terms of participants' self-reported demographics and their concentration or engagement in the task.

Further research is needed in which these issues are addressed. In particular, an assessment of attitudes towards gender-fair language is desirable. Overall, the study gives further arguments to support gender fair language, both in our personal language use as well as in official communication guidelines.

References

1. Pusch, L. (1984). Das Deutsche als Männersprache. Frankfurt/Main: Suhrkamp.
2. 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.
3. Blake, C. & Klimmt, C. (2010). Geschlechtergerechte Formulierungen in Nachrichtentexten. *Publizistik, 55*, 289–304.
4. Von Stockhausen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the represnetation of human referents. *Experimental Psychology, 57*, 367-375.
5. Ferstl, E. C., & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. Oostendorp &. S. Goldman (Hrsg.), *The construction of mental models during reading*, (pp. 247 - 277). Mahwah, NJ: Lawrence Erlbaum Associates.
6. Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13(2),* 168– 171.
7. Misersky, J., Gygax, P., Canal, P., … & Sczesny, S. (2014). Beauticians are rarely male: Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods, 46,* 841-871.

Time pressure effects on semantic speech-gesture coordination

Kirsten Bergmann and Stefan Kopp

Bielefeld University, Faculty of Technology, CITEC,
P.O. Box 100 131, D-33501 Bielefeld
{kirsten.bergmann,stefan.kopp}@uni-bielefeld.de
WWW home page: http://www.techfak.uni-bielefeld.de/ags/soa/

1 Introduction

Depending on how much information gestures convey beyond the meaning of the words they accompany, gestures can be classified along a continuum of (non-) redundancy. Empirical evidence shows that the way how speakers actually coordinate speech and gesture semantics is affected by cognitive factors like speech production problems, conceptualization load, or inter-individual differences in terms of verbal or spatial skills (cf. [1]). In the literature, different models of speech and gesture production have been proposed (e.g., [2]), but a concrete picture of the cognitive processes underling semantic coordination is still missing and many questions about the detailed mechanisms remain open.

A promising approach to explicate and test hypotheses are cognitive models that allow for computational simulation. In previous work, we developed a computational model to simulate semantic speech-gesture coordination in terms of the underlying cognitive processes [1]. Integrated into an overall speech-gesture production framework, a multimodal working memory serves as the central component in this model. In line with theoretical models, it comprises a symbolic-propositional representation for speech-, and a visuo-spatial one for gesture production. As an interface between these modality-specific representations, supramodal concepts are implemented which link visuo-spatial properties to corresponding propositional denotation. Cognitive processes operating upon the memory structures are modeled in terms of dynamic activation spreading principles.

We quantified our modeling results in simulation experiments in which we manipulated the available time (in terms of memory update cycles) before the model has to come up with a sentence and a gesture. We analyzed the resulting multimodal utterances with respect to semantic coordination: Non-redundant gestures showed to be dominant in those runs with stricter temporal limitations, while redundant ones become more likely when the time available was increased. In the present paper we present first results from an empirical validation study which aimed to provide adequate data from human speakers against which model predictions can be evaluated.

2 Study

To validate the predictions of our model, we designed a controlled experiment in which we empirically tested the impact of time pressure for speaking on the semantic coordination of speech and gesture. We manipulated the time available for speakers to give particular object descriptions: In a *high* time pressure condition the available time to give a particular object description was 15 seconds, while in a *low* time pressure condition the available time to provide a particular object description was 30 seconds. A total of 42 participants, aged from 19 to 41 years, took part in the study. 26 participants were female and 16 were male. All of them were recruited at Bielefeld University and received 2 Euros for participating.

Procedure The experiment was conducted in two consecutive phases. In the first phase, participants were provided with five stimulus pictures of buildings from a virtual town, together with their labels (same landmark buildings as in the SaGA corpus stimulus; cf [3]). They were left alone to look at each of the pictures as long as they thought necessary to memorize it. Subsequently, participants described the five buildings to a confederate who gave feedback, e.g. by nodding, or asked intermediate questions if necessary, but did not use any gestures herself to avoid priming participants' gestural behavior. The order of descriptions was randomized. Each description trial was preceded by a 5sec countdown. Thereafter, participants saw the label of the building to be described and a diminishing bar displaying how much time was left for the description. 2.5 seconds before the time limit was reached, a short sound signal was played, followed by a longer tone when the time limit was finally reached. At this point the description had to be stopped immediately. To get familiar with time-controlled descriptions and the procedure, participants – before starting the first experimental description – underwent a test trial in which there was nothing to describe, but everything else was as in the experimental trials.

Data Coding Both the speech and gestures were analyzed with respect to the semantic information they represented, based on an established micro-analytic coding method using a range of semantic features (e.g. [4]). The set of semantic features included in this analysis was considered to capture the kind of semantic information contained in the object description data. Our analysis focused on the amount of information represented, regardless of whether the information was complementary or redundant with regard to the information in the respective other modality. Hence, verbal utterances, as well as representational gestures were analyzed for the semantic information they contained, based on the following semantic categories: amount, entity, relative position, shape, and size. When the same semantic feature was covered by both speech and gesture, we considered this a redundant one, when gestures encoded semantic features not covered by the accompanying words, these were considered to be non-redundant.

As annotation-based data might be problematic due to subjective judgements of the coders, 28% of the data has been annotated independently by two annotators to investigate the degree of reliability. Cohen's Kappa was calculated, as a metric to evaluate data on a nominal scale. Substantial agreement was reached with Kappa values of κ=0.76 for gesture semantics and κ=0.86 for speech semantics.

Results Participants gave a total of 210 object descriptions in which they produced a total of 871 gestures. For several reasons we excluded some descriptions from further analyses. First, descriptions were not taken into account in which participants used 22 seconds or less of their time available in the 30sec condition (i.e., the description was closer to the 15sec condition than to the 30sec condition). Second, descriptions were excluded when participants were obviously not aware of the time pressure put on them which became apparent by utterances like "Oh, I did not notice that description time was so short" at the end of the very descriptions. Finally, descriptions were excluded in which participants made meta comments, e.g. when they mixed up the different objects to be described and got aware of this during their description. In the end, at total of 645 gestures remained for further analyses (261 in the 15sec condition, 384 in the 30sec condition).

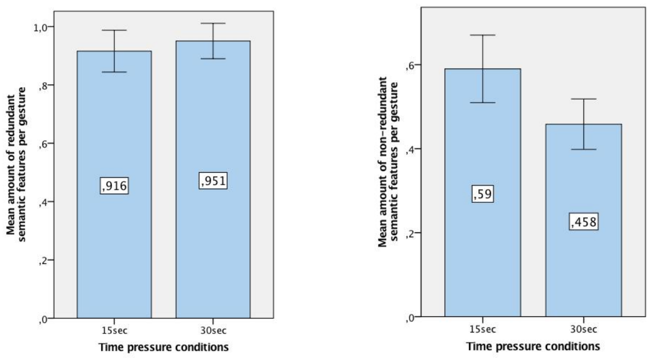


Fig. 1. Mean amount of redundant (left) and non-redundant (right) semantic features per gesture. Error bars represent 95% confidence interval.

A one-way ANOVA was conducted to compare the effect of time pressure on semantic speech-gesture coordination. Results are visualized in Figure 1. There was a significant main effect of time pressure on the amount of non-redundant semantic features ($F(1,\ 643) = 6.92$, $p = 0.009$). The mean amount of non-redundant semantic features in gestures of the 15sec condition was increased as compared to gestures produced in the 30sec condition. For the amount of redundant semantic features, there was no main effect ($F(1,\ 643) = 0.53$, $p = 0.469$).

3 Discussion

The study presented here aimed to evaluate of a cognitive model of semantic speech-gesture coordination in comparison with empirical data [1]. The model predicts (i) non-redundant gestures to be dominant under stricter temporal limitations, while (ii) redundant ones become more likely when time available is increased. First empirical results support the first prediction, whereas redundant gestures showed to occur at rather equal rates under high vs. low time pressure. That is, first of all, the empirical data supports the notion of time pressure affecting semantic speech-gesture coordination. However, supportive evidence is only partial. Several explanations are conceivable for this finding. For instance, it might be that the operationalization of time pressure, as employed in the experiment, is not fully equivalent to that of the modeling approach. In particular, the empirical investigation was not able to put time pressure on the planning process of any single utterance. As participants had to give a longer description, respectively, they had the chance to compensate for temporal constraints in several ways, e.g. by reducing the total amount of information in their communicative behavior, by planning shorter/longer increments, or by compensating for longer planning time by increasing the speech rate. Moreover, the presence of an interlocutor as in our empirical study, a variable not yet considered in the model, might affect the way speakers plan and realize their multimodal communicative behavior. Or, the assumptions underlying our model of speech-gesture coordination processes or its parametrization might need some adjustment to match empirical data. More detailed analyses of the empirical data are underway to elucidate these issues. In particular, high vs. low time pressure data will be analyzed with respect to further variables such as speech rate, amount of (filled) pauses, temporal speech-gesture coordination etc.

Acknowledgements This research was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

1. S. Kopp, K. Bergmann, and S. Kahl, "A spreading-activation model of the semantic coordination of speech and gesture," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society (CogSci 2013)*, 2013, pp. 823–828.
2. S. Kita and A. Özyürek, "What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking," *Journal of Memory and Language*, vol. 48, pp. 16–32, 2003.
3. A. Lücking, K. Bergmann, F. Hahn, S. Kopp, and H. Rieser, "Data-based analysis of speech and gesture: The Bielefeld speech and gesture alignment corpus (SaGA) and its applications." *Journal on Multimodal User Interfaces*, vol. 7, no. 1-2, pp. 5–18, 2013.
4. J. Holler and G. Beattie, "How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process?" *Semiotica*, vol. 146/1, pp. 81–116, 2003.

Wall Distance as a Cue in Human Place Recognition

Marc Halfmann[1,2], Viktoria Prozmann[1], Nina Walker[1], and Hanspeter A. Mallot[1]

[1] Dept. of Biology, University of Tübingen, Tübingen, Germany,
hanspeter.mallot@uni-tuebingen.de
[2] Leibniz-Institut für Wissensmedien, Tübingen, Germany

Abstract. Place recognition from visual cues involves the standard processes of early vision, including among others the detection of image features and depth, the understanding of scenes, and the recognition of objects. Here we use stereoscopic dynamic random dots to study the role of pure depth information in place recognition. Results indicate that place recognition can be based on pure depth information and is not substantially improved by cues from other visual sub-modalities such as texture or localized objects (room corners).

1 Introduction

The recognition of places is generally thought to rely on a combination of landmark cues visible from the target place and spatial context such as traveled distances from neighboring places (e.g., O'Keefe & Nadel, 1978). For the landmark component, various types of information can be extracted from the visual input and have been shown to play a role in place recognition. These types include barely processed "snapshots" (for review, see Gillner et al., 2008) as well as visual information requiring higher amounts of image processing such as landmark configurations (Mallot & Lancier, 2016), room geometry and three-dimensional spatial layout (Cheng et al., 2013; Epstein et al., 2008), or identified landmark objects (Janzen & van Turenout, 2004). Visual depth, i.e. the perceived distance to objects of the surrounding scene, is relevant for a number of these cues, especially if indoor-environments are considered. Here we use psychophysical approaches from the study of early visual processes (stereopsis, motion parallax) to investigate the role of perceived depth in place recognition.

2 Methods

Subjects and Procedure. 40 students from the University of Tübingen passed a simple test for stereo vision and participated in this study. The experiments were carried out in a virtual environment simulating a kite-shaped room with edged or rounded corners. In the "return-to-cued-location task" (Gillner et al., 2008), participants were placed at one of three goal locations in the kite-shaped

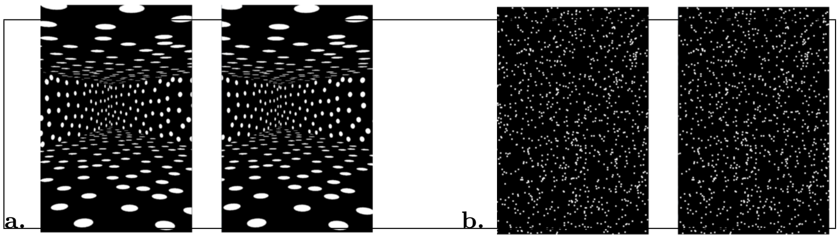


Fig. 1. View of the kite-shaped room arranged for free stereoscopic viewing using crossed fusion. **a.** Texture condition. **b.** Dot condition (sample frame of the dynamic random dot display). Both stereograms show the room with edged corners.

room. In the following inspection phase subjects studied the local appearance of the room by looking around and performing small translational movements. They were then set back to a start position and used a joy-stick to return to the goal. After indicating goal recognition by the button hit, subjects were moved to the correct goal position, and the next trial started from there. In all, twelve decisions were recorded per subject and condition, i.e. two cycles of all six possible transitions between the three goal locations. In the results reported here, the virtual environment was presented with an Oculus-Rift stereoscopic head-mounted display (HMD), but controls with a mirror stereoscope and monocular viewing were also performed. In addition to the stereo disparities presented on the stereoscope, the HMD setup provided a higher level of immersion including closed-loop movements of the head and body that might lead to better perception of structure-from-motion. We thus hypothesized that using a HMD would increase participants' place recognition performance as well as decreasing the response time needed.

Stimuli and Conditions. Two factors, "visual cues" and "room shape", were varied in a full factorial design. In the *cue-condition "texture"*, rooms were defined by a texture of large spots (about 10 cm diameter in the virtual environment) pasted to the room walls, floor, and ceiling as a wallpaper. This texture provided stereo disparity, motion parallax upon observer motion, texture gradients and information about room corners (Fig. 1a). In the *cue-condition "dots"*, surfaces were defined by dynamic random dots with limited lifetime varying between 100 and 200 ms. Outdated dots or dots leaving the field of view were continuously replaced so that the dot distribution on the screen was kept uniform. The dots provided stereo disparities, a small amount of motion parallax (during dot lifetime), but no texture gradients. Room corners might have been inferred from the depth information, but not from the dot distribution itself (Fig. 1b). The cue conditions were performed in a blocked, within-subject design (dots condition first). We used two *shape-conditions* "edged", and "rounded", as shown in Fig. 2 (between subjects factor). With these we wanted to test the hypothesis that the corners of the room provide conspicuous landmarks that might be important for place recognition.

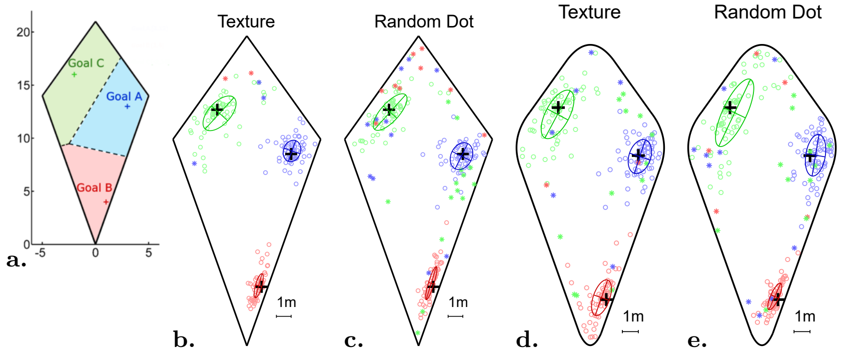


Fig. 2. Scatter plots of the decision points from 240 decisions (20 subjects × 12 decisions per subject). **a.** Layout of the kite-shaped room with goal locations A, B, C, and nearest-neighbor cells. Dimensions in meters. **b., c.** Edged corner room, **d., e.** rounded corner room. Dot colors indicate goal positions A, B, C. Tokens indicate: + true goal location, ∘ decision points within goal region ("correct decision"). ∗ decision point outside goal region ("qualitative error"). Error ellipses are calculated over the within-region decisions only and reflect one standard deviation.

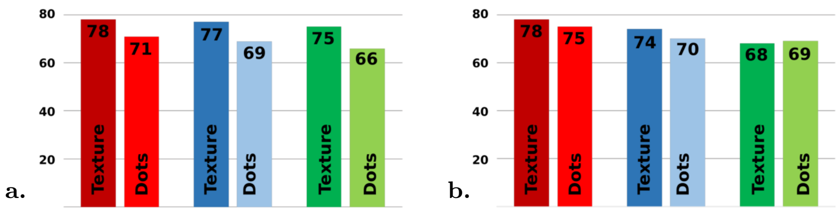


Fig. 3. Absolute numbers of correct decisions (decisions inside goal region) out of a total of 80 decisions per target (accumulated over all subjects). **a.** Edged corner room, **b.** rounded corner room. Colors indicate goal locations A, B, C. Performance above chance level defined by the relative area of Voronoi cell is highly significant for all cases.

3 Results

Fig. 2 shows the decision points in the four conditions, accummulated over all subjects. Decision points scatter about the goal positions with a moderate variance, and variance is not substantially different in the four conditions. We also find a fair number of "qualitative errors" in which the subjects choose a place closer to one of the non-goals than to the current goal. The respective nearest-neighbor cells (Voronoi tesselation around goal points) are also indicated in Fig. 2a. These errors are equivalent to the "rotation errors" discussed in the geometric-module literature (see Cheng et al. 2013 for review). Fig. 3 shows the number of correct decisions for the various conditions, again accumulated over all subjects. If subjects would ignore the visual information, the chance level for chosing a decision point in the correct Voronoi cell would be about 33% compared to an average recorded performance rate of about 91% shown in Fig. 3.

The difference from chance level is highly significant in all cases. No significant differences between conditions were found.

A comparison with the stereoscopic and monocular viewing conditions (data not presented in this paper) shows similar results. Performance is well above chance even for the monocular condition, albeit slightly poorer than in the HMD-data reported here.

4 Discussion

The results indicate that subjects can use pure depth information as is provided by dynamic random dots to recognize places in a room. Additional texture cues providing more reliable depth information seem to lead to some improvement, which, however, is not statistically significant. This is even more surprising since texture cues provide still another cue for place recognition, i.e. snapshot matching. Indeed, since the texture was "painted to the wall", the subjects might have tried to remember the pattern of black and white wall texture appearing at each goal location and try to match it to their memory when they return. If they did use this strategy, it did not lead to a substantial improvement in performance. The corners of the room do not seem to play an important role in self-localization, indicating that subjects rely more on the distances to walls than to the corners. Overall, the results fit nicely to the idea that places are represented by a local map of the environment which is updated as the subject moves around (Byrne et al., 2007, Loomis et al.; 2014, Röhrich et al., 2014).

References

1. Byrne, P., Becker, S., Burgess, N. (2007). Remembering the past and imagining the future: A neural model of spatial memory and imagery. *Psych Rev*, 114:340 – 375.
2. Cheng, K., Huttenlocher, J., Newcombe, N. S. (2013). 25 years of research on the use of geometry in spatial reorientation: a current theoretical perspective. *Psychon Bull Rev*, 20:1033 – 1054.
3. Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends Cogn Sci*, 12:388 – 396.
4. Gillner, S., Weiß, A. M., Mallot, H. (2008). Visual place recognition and homing in the absence of feature-based landmark information. *Cognition*, 109:105 – 122.
5. Janzen, G., van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7:673 – 677.
6. Loomis, J. M., Klatzky, R. L., Giudice, N. A. (2013). Representing 3D space in working memory In Lacey, S., Lawson, R., eds, *Multisensory Imagery: Theory and Applications*, pp. 131 – 156. Springer, New York.
7. Mallot H.A., Lancier, S. (2016) A maximum-likelihood approach to place recognition from distant landmarks. *This volume.*
8. O'Keefe, J., Nadel, L. (1978). *The hippocampus as a cognitive map.* Clarendon, Oxford, England.
9. Röhrich, W., Hardiess, G., Mallot, H. A. (2014). View-based organization and interplay of spatial working and longterm memories. *PlosONE*, 9(11):e112793.

A Maximum-Likelihood Approach to Place Recognition from Distant Landmarks

Hanspeter A. Mallot[1] and Stephan Lancier[1]

Dept. of Biology, University of Tübingen, Tübingen, Germany,
hanspeter.mallot@uni-tuebingen.de

Abstract. We present a probabilistic model of place recognition from a configuration of distant landmarks surrounding a goal. The model assumes that landmark positions are perceived with hyperbolic distance compression and added noise, depending on current observer position. Position-dependent recognition rate is modeled as the likelihood of perceiving the expected (stored) landmark configuration from each position. The model reproduces key features of experimental results including a systematic localization bias towards the most distant landmark, the shape and orientation of the error ellipses, and effects of approach direction. We conclude that place recognition is based on a comparison between a place code (landmark distance and angles) and a working memory of surrounding space suffering from systematic depth distortions and distance-dependent drop in resolution.

1 Introduction

Place recognition, like any other recognition task, has to be based on a comparison between a reference coded in memory and the current sensory input. This reference could be a raw snapshot of the scene taken from the goal location as has been suggested for honey-bee and ant navigation; for snapshot use in humans, see Gillner et al. (2008). In experiments with configurations of isolated landmark objects (Waller et al. 2002, Pickup et al. 2013), the role of depth information in place recognition has been demonstrated. Here we review experimental data from Lancier (2016) and present a quantitative, maximum likelihood model of place recognition involving a memory-code for place which is based on landmark distance and bearing.

2 Model constraints

The accuracy of the place recognition in an open environment comprising four distant, distinguishable landmarks was studied in a behavioral experiment with human subjects navigating a virtual environment. The environment included a +-shaped bridge crossing a pond and four colored spheres hoovering in mid-air above the pond, one in each quadrant defined by the bridge arms. Subjects started at one bridge entry and had to find a goal that involved either a left or a right turn at the bridge center ("decision point"). All possible starting

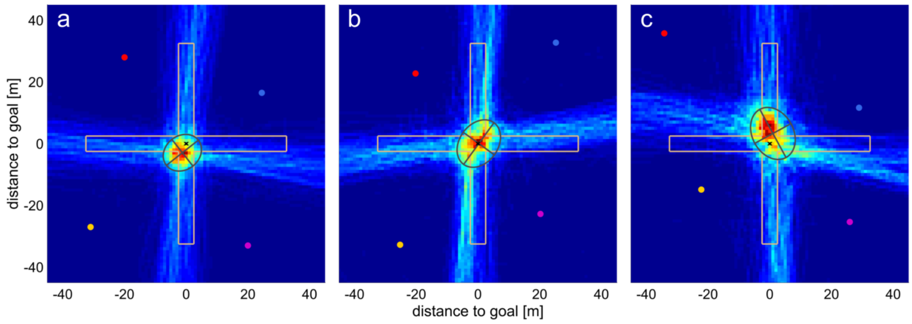


Fig. 1. Position choices for three landmark configurations. The landmarks are shown with their actual position and color. **a.** Standard configuration (20 subjects, 954 decisions), **b.** Parallelogram configuration (16 subjects, 761 decision), **c.** Peaked configuration (16 subjects 754) decision. Mean deviation from the bridge center was significant for the standard and peaked conditions (Hotelling's T-Square test). The orientation towards the most distant landmarks in a and c was tested with a circular V-test over the preferences of the individual subjects and was also significant. See Lancier (2016).

points and turn directions were used. In the test phase, bridge, pond, and goals were rendered invisible by simulated ground fog and the subjects were asked to navigate to the now invisible center of the bridge and indicate place recognition by button hit. This performance was based essentially on the four landmarks which remained visible at all times. In order to prevent subjects from using path integration, the starting points at each of the four bridge entries were varied using a random positional scatter. Experimental results are summarized in Figure 1 (Lancier 2016). For the model, the following constraints can be derived:

1. Decision points show both a systematic bias and a statistical error. The systematic bias as well as the major axis of the error ellipses point roughly in the direction of the most distant landmark (Fig.1a-c).
2. If a point-symmetric configuration of landmarks is used, the systematic bias goes away (Fig.1b).
3. If the landmark sizes, and therefore the perceived landmark distances, are manipulated between training and test session, decision points are shifted towards downscaled landmarks and away from the upscaled ones (data not shown). I.e. subjects try to adjust remembered and perceived distances.

3 Model

In a world coordinate system centered around the target point (the center of the bridge), the landmark positions are denoted by $\boldsymbol{l}_i$, $i = 1, \ldots, 4$. Let $\boldsymbol{x}$ denote the current observer position. The true landmark vectors from the current observer position are $\boldsymbol{m}_i = \boldsymbol{l}_i - \boldsymbol{x}$. In order to model the systematic bias, we will need to assume that the actual perceived landmark distance is not veridical but hyperbolically compressed according the equation $\boldsymbol{\mu}_i = A\boldsymbol{m}_i/(A + \|\boldsymbol{m}_i\|)$ (Gilinsky

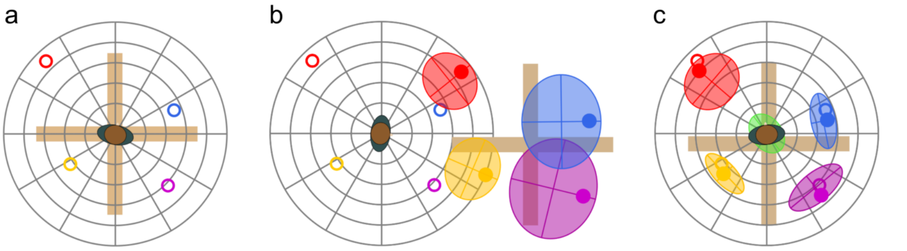


Fig. 2. Place recognition model. **a.** Layout with bridge and landmarks. **b.** Approaching observer with place code (open colored circles). Solid colored disks: true landmark positions; transparent ellipses: distribution of landmark measurement according to Eq. 1. **c.** Probabilistic match of place code and observed landmark positions.

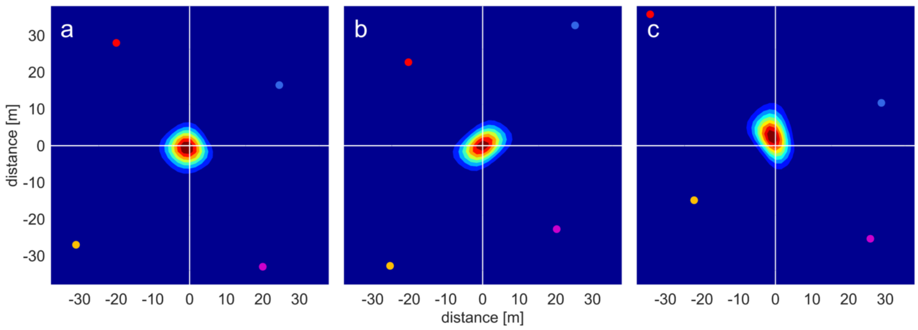


Fig. 3. Log likehood function (Eq. 3) for the three landmark configrations. Conventions as in Figure 1.

1951). A is a constant set to 100 m in our simulations. This compression does not affect the stored landmark position which is assumed to be derived from triangulation and spatial updating processes (Philbeck & Loomis 1997) and may therefore be assumed veridical. The stored place code is therefore given by the true landmark positions $\boldsymbol{l}_i$. Consider the probability of perceiving a landmark i at a position $\boldsymbol{m}_i$, given the current observation position is $\boldsymbol{x}$. This measurement $\boldsymbol{m}_i$ is given in Cartesian coordinates but comprises information about the egocentric perceived distance (with hyperbolic compression) and allocentric bearing, i.e. bearing with respect to some reference direction. The probability density function is assumed to be

$$p(\boldsymbol{m}_i|\boldsymbol{x}) = \phi(\boldsymbol{m}; \boldsymbol{\mu}_i(\boldsymbol{x}), \Sigma_i(\boldsymbol{x})), \tag{1}$$

i.e. the two-dimensional normal distribution with mean $\boldsymbol{\mu}_i$ and covariance matrix Σ_i. Note that both mean and covariance depend on the current observer position $\boldsymbol{x}$. For the mean, we have specified this dependence above. The covariance matrix Σ will have an eigenvector in the direction $(\boldsymbol{l}_i - \boldsymbol{x})^o$, i.e. the depth direction from the current view-point to the true landmark position, and a another one in the width direction,

$$\Sigma_i(\boldsymbol{x}) = R_i \begin{pmatrix} \lambda_d & 0 \\ 0 & \lambda_w \end{pmatrix} R_i^\top, \text{ where } R_i(\boldsymbol{x}) = \left[(\boldsymbol{l}_i - \boldsymbol{x})^\circ, (\boldsymbol{l}_i - \boldsymbol{x})^{\circ\perp}\right]. \quad (2)$$

The eigenvalues in the distance and width directions are assumed to scale with distance according to $\lambda_{id}(\boldsymbol{x}) = 0.0001\ \|\boldsymbol{l}_i - \boldsymbol{x}\|^4$ and $\lambda_{iw}(\boldsymbol{x}) = 0.1\ \|\boldsymbol{l}_i - \boldsymbol{x}\|^2$. Thus, the angular error of landmark position does not depend on viewing distance. For small distances, $\lambda_{iw} > \lambda_{id}$, as is necessary to model the shape of the experimental distributions. This may reflect the fact that inter-landmark angles have to be inferred from multiple views. The place code for the goal position $\boldsymbol{x} = 0$ will be $\{\boldsymbol{l}_i, i = 1, \ldots, 4\}$. The probability of measuring this place-code, given that the observer is actually at $\boldsymbol{x}$, is obtained by substituting $\boldsymbol{m} = \boldsymbol{l}_i$ in eq. 1 and taking the product over all four landmarks:

$$p(\boldsymbol{l}_1, ..., \boldsymbol{l}_4|\boldsymbol{x}) = \prod_{i=1}^{4} \phi(\boldsymbol{l}_i; \boldsymbol{\mu}_i(\boldsymbol{x}), \Sigma_i(\boldsymbol{x})). \quad (3)$$

The function $LL(\boldsymbol{x}) := \log p(\boldsymbol{l}_1, ..., \boldsymbol{l}_4|\boldsymbol{x})$ is plotted as the model prediction in Figure 3 for the three landmark configurations appearing also in Fig. 1.

4 Conclusion

Place recognition from distant landmarks is based on a comparison of two components, (i) a referential place code containing veridical landmark distances and inter-landmark angles, and (ii) a visual working memory of the complete surroundings with distance-dependent resolution and systematic depth compression. A simple model of these components is able to quantitatively predict the statistical distribution of decisions made by human subjects. Effects of approach direction can be modelled by increasing the variances of the less-seen landmarks.

References

1. Gilinsky, A. S. (1951). Perceived size and distance in visual space. *Psychological Review*, 58:460 – 482.
2. Gillner, S., Weiß, A. M., Mallot, H. (2008). Visual place recognition and homing in the absence of feature-based landmark information. *Cognition*, 109:105 – 122.
3. Lancier S., (2016) *Spatial Memories in Place Recognition*, PhD-Thesis, Faculty of Science, University of Tübingen
4. Philbeck, J. W., Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23:72 – 75.
5. Pickup, L. C., Fitzgibbon, A. W., Glennerster, A. (2013). Modelling human visual navigation using multi-view scene reconstruction. *Biological Cybernetics*, 107:449 – 464.
6. Waller, D., Loomis, J. M., Golledge, R. G., Beall, A. C. (2002) Place learning in humans: The role of distance and direction information. *Spatial Cognition and Computation*, 2:333 – 354

Self-Localization Accuracy and Spatial Ambiguity of Humans and Robots in a Complex Building

Rul von Stülpnagel[1*], Vincent Langenfeld[1*], and Christoph Hölscher[2]

[1] Freiburg University, Germany
rul.von.stuelpnagel@cognition.uni-freiburg.de;
langenfv@informatik.uni-freiburg.de
[2] ETH Zürich, Switzerland
christoph.hoelscher@gess.ethz.ch
* Equal contribution

Abstract. We investigated differences and commonalities of robots' and humans' self-localization abilities in a complex building without landmarks as well as with landmarks, optimized for humans and robots, respectively. Our findings exemplify the fundamental differences between robot and human processing of spatial information. This research outlines the challenges for all scenarios that encompass robot-human cooperation in regard to spatial orientation and navigation.

Keywords: Self-localization, spatial ambiguity, robot-human comparison

1 Introduction

Architectonically complex, but visually ambiguous buildings (e.g., conference buildings or museums) represent a challenge for both robot and human orientation. However, the sensor apparatus of humans and robots differs as much as their means to process spatial information. Humans excel at identifying and recognizing distinctive features in their environment. The comparatively unlimited memory capacities of robots allow for a precise matching of the current vista space with all other known vista spaces. The growing number of potential interactions (e.g., search & rescue missions in a partly collapsed building) requires a better understanding of commonalities and differences of humans and robots in regard to spatial orientation. At the core of spatial orientation, successful self-localization is a crucial issue. This research aims at excluding several aspects involved in navigation and orientation in order to allow for a systematic comparison of self-localization accuracy at isolated locations within a familiar environment between humans and robots.

There have been converging approaches to reduce spatial ambiguity during self-localization through ad-hoc installment of landmarks. The deployment of RFID chips as landmarks fostered the mapping abilities of robots [1], [2]. Humans with access to individually placed or preplaced landmarks showed superior wayfinding performance in a virtual environment as compared to participants without landmarks [3]. Another study did not find such advantage, but reported strong consistencies in human landmark placement [4]. Thus, a second goal of this research concerns the question how and where landmarks must be placed to reduce spatial ambiguity of robots and humans.

2 Method

We tackled our research questions in a complex building (see Fig. 1). We applied a grid of 1×1m cells on this layout, resulting in 3,101 cells total for all following computations. In *Stage 1*, the building contained no landmarks other than provided by the layout itself. We applied the uniqueness measure introduced by Meyer-Delius et al. [2] to quantify robot self-localization accuracy. The uniqueness (robot *U*) of a pose *x* was defined by the probability of obtaining the same sensor observation (simulating a 2D LRF scan) as in *x* from any other pose in the environment, averaged over all other poses. Human self-location accuracy was evaluated by selecting 100 cells based on maxima and minima of several criteria (e.g., the number of visible cells, robot U, etc.). Fifteen undergraduates were familiarized with the layout before studying first-person views of the 100 locations in a desktop virtual environment. After studying each location, they estimated their position on an empty floorplan. An estimation was counted as successful if the position was within a radius of 15m (reflecting the limit of clearly distinguishability on the computer screen) and within line of sight of the original position. We collapsed the data across participants and computed the means for localization accuracy per location. Figure 1 depicts robot and human self-localization accuracy of all stages.

For *Stage 2* (and based on [1] and [4]), we computed which five cells of the environment had to contain a landmark to minimize robot ambiguity as indicated by mean robot *U*, based on the 100 locations selected in Stage 1 only and under the assumption that a detected landmark enables perfect self-localization. Sixteen new undergraduates studied the same locations as in Stage 1 and estimated their position. Landmarks were visible both from the egocentric and the allocentric perspective.

For *Stage 3*, we identified five landmark positions optimized for human self-localization, based on the assumption that a visible landmark enables perfect self-localization and the probability of successful self-localization at each location in Stage 1. The evaluation of robot and human self-localization was equivalent to Stage 2, with the landmark positions and the data collection of sixteen new undergraduates being the only difference.

3 Results

We scaled the values of robot U for all cells to range between 0 and 1 (with 0 and 1 representing the cells with the lowest and highest uniqueness value, respectively, across all experiment stages). As expected, we found the poorest robot U for ‘no landmarks’ $(M = .55,\ SD = .31)$, an enhanced robot U for ‘optimized for humans’ ($M = .66$, $SD = .29$), and the highest robot U for ‘optimized for robots’ $(M = .75,\ SD = .25)$. We refrained from a statistical analysis as robot U cannot be treated as a continuous variable with a normal distribution. Similarly, we found the poorest human localization accuracy for ‘no landmarks’ $(M = .57,\ SD = .20)$. Contrasting robot self-localization, we found a significantly increased accuracy with landmarks for optimized for robots $(M = .68,\ SD = .22)$, $t(99) = 5.82,\ p < .001$, but even better performance with landmarks optimized for humans $(M = .80,\ SD = .18)$, $t(99) = 4.99,\ p < .001$.

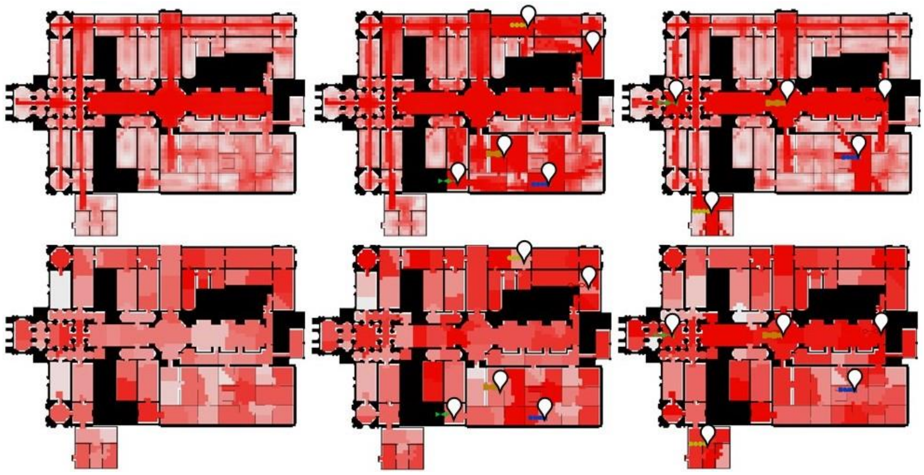


Fig. 1. Self-localization accuracy of robots (top row) and humans (bottom row). Darker red indicates higher self-localization accuracy in *Stage 1* 'no landmarks' (left), in *Stage 2* 'landmarks optimized for robots' (center), and in *Stage 3* 'landmarks optimized for humans' (right). Landmarks (white markers) are portrayed uniformly rather for reasons of clarity, but were individually identifiable for both robots and humans.

4 Discussion

We compared the self-localization accuracy of robots and humans in a familiar (virtual) building. Given the unusual and cognitively demanding task of transferring the egocentric study perspective to the allocentric localization perspective, humans performed surprisingly well even without landmarks. Their performance increased when landmarks were available. However, the significantly higher localization accuracy supported by landmarks optimized for humans as compared to those optimized for robots implies that the "where" of these landmarks must be aligned to human spatial cognition (see [4]). The data indicate a reversed pattern for robots' self-localization accuracy, thus implying that robot and human spatial ambiguity results from very different reasons. Robot ambiguity was evaluated by comparing the similarity of a (simulated 2D LRF) laser-scan to all other possible scans. Provided with perfect "memory", robot self-localization benefits from increasing complexity of the vista space. In contrast, it is not clear yet which spatial and environmental factors enhance and inhibit human self-localization performance. Preliminary analyses indicate that human self-localization accuracy did not depend on spatial properties of a location such as size, uniqueness, or jaggedness of the visible space. We also found no evidence that human performance is determined by distinctive local features (e.g., the style and number of doors in a room). We are currently investigating whether the shape of the investigated room as well as the structural composition of the surrounding rooms (i.e., similar to a jigsaw puzzle) is a better predictor of human self-localization accuracy [5].

The differences between spatial ambiguity of humans and robots become also clearly visible in the distribution of landmarks optimized for robots and humans, respectively. Landmarks optimized for robots are located in segregated areas where high ambiguity was measured, as robots are most unlikely to make errors in the uniquely formed central areas of the building. In contrast, landmarks optimized for humans were tightly linked to their visibility and integration in the general layout of the building (i.e., its central areas), thus indicating that humans take the global structure of the environment into account during self-localization. Landmarks served to eliminate residual ambiguity for a large part of the building (e.g., the main hall) rather than to disambiguate specific areas.

Our findings shed first light onto the challenges that need to be addressed for any scenario that encompasses robot-human cooperation in regard to spatial orientation and navigation. Robot orientation may for example be impeded by architectural designs geared towards human needs, with a small number of highly distinctive and central locations, but a large number of uniform functional rooms. Future research should thus take the underlying structure of different building types into account. Furthermore, we deliberately limited the current approach to static self-localization in order to increase the comparability of human and robot performance. Extending our approach to a scenario featuring locomotion represents a worthwhile challenge. Insights gained from such comparisons become more and more relevant with the increasing integration of robots into human lives.

5 References

[1] M. Beinhofer, H. Kretzschmar, and W. Burgard, “Deploying artificial landmarks to foster data association in simultaneous localization and mapping,” in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, 2013, pp. 5235–5240.

[2] D. Meyer-Delius, M. Beinhofer, A. Kleiner, and W. Burgard, “Using artificial landmarks to reduce the ambiguity in the environment of a mobile robot,” in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, 2011, pp. 5173–5178.

[3] D. Cliburn, T. Winlock, S. Rilea, and M. Van Donsel, “Dynamic landmark placement as a navigation aid in virtual worlds,” in *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, 2007, pp. 211–214.

[4] R. Von Stülpnagel, S. Kuliga, S. J. Büchner, and C. Hölscher, “Supra-individual consistencies in navigator-driven landmark placement for spatial learning,” in *Proceedings of the 36th annual conference of the cognitive science society. Cognitive Science Society, Austin, TX*, 2014, pp. 1706–1711.

[5] B. J. Stankiewicz and A. A. Kalia, “Acquisition of structural versus object landmark knowledge.,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 33, no. 2, pp. 378–390, 2007.

Modelling different strategies in mental rotation

Alexander Lotz* & Nele Russwinkel**

Cognitive Modeling in Dynamic Human-Machine Systems, Technische Universität Berlin, Germany. Tel: +49 (0)30-314-29631
* e-mail: alexander.lotz@tu-berlin.de
** e-mail: nele.russwinkel@tu-berlin.de

Abstract. This study aims to analyse the linear correlation (relationship between reaction times and angular disparity) presented in many past publications regarding mental rotation (e.g. Shepard & Metzler 1971), by modelling the mental rotation task in ACT-R (Anderson et al. 2004). Furthermore we wish to distinguish whether the mental rotation task depends on *process-* or *instance-based* theories. We studied how the introduction of a new object and new rotation axes after learning the mental rotation task affected the reaction times of subjects. Results show that linear mental rotation rates could not be found in our data and new objects increased reaction times. This leads us to conclude that pure *process-based* theories do not solely apply and proposes that declarative knowledge is needed for mental rotation.

Keywords: ACT-R, Mental Rotation, Strategy Choice, Spatial Competence

1 Introduction

When practising certain skills two primary mechanisms are assumed to explain learning effects, i.e. *process-based* and *instance-based* theories (Logan 1988; Heil et al. 1998). These general theories apply to a large variety of tasks that can be learned through repetition, including tasks relying on spatial competence. Spatial competence, thereby, is a key skill humans need to possess when planning manipulations of objects. One such elementary spatial task is mental rotation. Many experimental results have been published on mental rotation with contradictory conclusions on learning effects, splitting into *process-based* and *instance-based* explanations (*process-based*: Wallace & Hofelich 1992; *instance-based*: Tarr & Pinker 1989; Kail & Park 1990).
We aim to model mental rotation with a cognitive architecture, i.e. Adaptive Control of Thought-Rational (*ACT-R*; Anderson et al. 2004), to predict user behaviour in spatial tasks. These tasks can be as rudimentary as mentally planning an object manipulation or complex interaction with interfaces in common human-machine-interactions. While experimental results on mental rotation cluster into the two abovementioned groups, it is important to understand that these theories rely on fundamentally different forms of knowledge, i.e. procedural and declarative knowledge. As it is our goal to correctly identify and model mechanisms that influence users' ability, e.g. control machinery with interfaces, these two forms of knowledge are an important differentiating factors and cannot be handled identically when modelling cognition.

2 Experiment

A total of 36 volunteers (17 female, 19 male) were recruited via a participant database of the Technische Univeristät Berlin. The sample primarily included students of the university between 18 and 30 years of age (mean= 24.9). A 2x2 between-subject design was chosen with factors a) number of objects and b) number of rotation axes. The independent variables were: object(s) which were rotated about 18 angular disparities, in two rotation directions around environmental axes (horizontal and/or vertical) with answers either being ("same") or ("different"). If subjects were assigned to the single object or single axis condition, trials were presented twice (randomly distributed) to obtain the fixed sum of 320 trials. Additionally, in the final block a new object was introduced that was rotated about four angular disparities (10°, 60°, 110° and 160°) in two rotation directions around identical two axes with answers also either being ("same") or ("different"). The dependent variables were the reaction time and the answer given.
This set-up was divided into four blocks with blocks one to three having a length of 72 trials, the final block having a length 104 trials, due to the introduction of the new object. Objects presented were chosen from a set size of four, the additional object of the final block also being chosen from a separate set with a set size of four. All trials were presented on a 22" LCD monitor and followed the experimental design of Shepard & Metzler (1971). Participants had normal or corrected-to-normal vision. All subjects received either course credit for their respective studies or 10 Euros.
For this study we hypothesise, among others (to be presented in the talk), that:

- *Operation hypothesis*: Rotation times will decrease of the course of the experiment, differentiation between *process-based* and *instance-based* theory is not possible.
- *Object hypothesis*: Rotation times for new objects introduced in the fourth block are longer than for the primary object, i.e. *instance-based* theory.

3 Model of Mental Rotation

As a basis for this work, previous modelling results are used as a reference (see Fig. 1; Lotz & Russwinkel (submitted)) and the implemented strategy of said ACT-R model is redefined. Prior experimental and simulation results, that were conducted with the angular disparities of (10°, 60°, 110° and 160°), did not confirm the linear correlation by Shepard & Metzler (1971). Therefore, the new model will choose between three different cases when solving the mental rotation task to allow modelling of non-linear mental rotation effects. It's to be pointed out that the abovementioned angular disparities are identical to those introduced in the fourth block of the experiment of this paper. The reasoning being that we can compare reaction times to previous results and evaluate whether reaction times and errors are in line with previously obtained results, giving an indication which of the theories applies to this scenario.
First, trials with small angular disparities are solved on the basis of the *instance-based* theory. When comparing the two images of a trial, representations of the images are

created (reference representation and rotation representation) and subsequently compared. If the rotation of the two representations is below a threshold of 20° (this value is an estimate, which has to be validated by empirical results of this study), we assume that no mental rotation is needed to solve the task. As representations of the objects are repeatedly created during the course of the experiment, calls to the declarative memory become faster, resulting in quicker reaction times.

Second, a specific strategy for rotations of 180° about an axis is implemented. We assume that this distinct inverse representation can be obtained without mental rotation being applied. Learning in this specific case is also due to representation instances of the objects being created, leading to lower declarative retrieval times.

Third, all other trials with angular disparities between 20° and 180° are modelled by a mental rotation process. This process combines the *process-based* and *instance-based* theories by learning the axis for the mental rotation (process-based) and, similarly to the previous two strategies, representations increasing activation in the declarative memory (instance-based), this strategy is already implemented (see Fig. 1).

By modelling the empirical data we aim to validate different cases, which may apply, when regarding mental rotation as possible explanations for non-linear rotation rates. Additionally, we wish to investigate the hypothesised threshold of 20°.

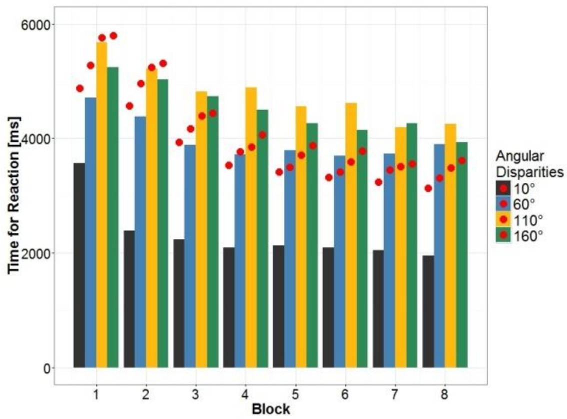


Fig. 1. Reaction times plotted for each block and angular disparity. Results displaying the model fit (red dots) compared to experimental data (Lotz & Russwinkel (submitted)). Modelling results are limited to the third case as described below. The modeling results show, that this case does not suffice to accurately model the empirical data.

4 Results

The following results were analysed with regard to subjects' reaction times. Figure 2 displays the reaction times with regard to blocks (course of the experiment) and the angular disparity. In all blocks complete linearity, as proposed by Shepard & Metzler (1971), is not present. Our *operation hypothesis* was subject to an ANOVA which led

to significant results $F(1,35)= 431$, $p<0.001$. An ANOVA conducted on the *object hypothesis,* delivered statistically significant results $F(1,35)$, $p=0.03$. Further hypotheses differentiating between-subjects groups and the 20° threshold will be discussed in the talk.
Furthermore, the simulation results of the model will be compared to the experimental data. The strategies integrated into the model that depend on theories, i.e. *process*- and *instance-based* theories of learning, promise a further step towards successfully modelling mental rotation as a key skill of human spatial competence.

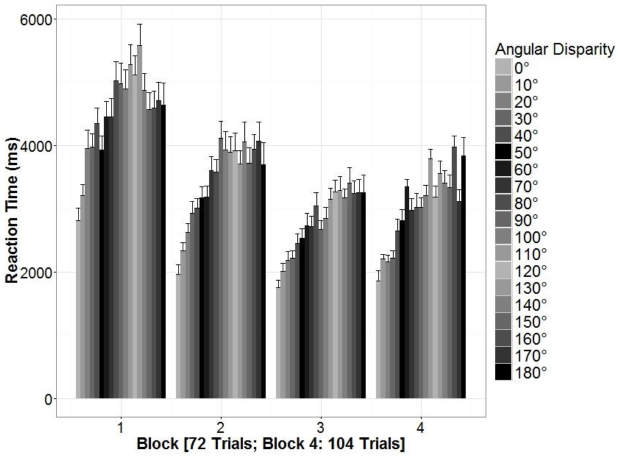


Fig. 2. Reaction times displayed for each experimental block and the angular disparities. A clear reduction of reaction times over the course of the experiment is shown, which is statistically significant. For angles greater than 90° reaction times converge. In the final block (block 4) no reaction time improvement is achieved and angular disparities of the new object presented (10°, 60°, 110° and 160°) increase.

References

1. Anderson, J., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C. and Y. Qin (2004). An Integrated Theory of Mind. *Psychological Review*. **111** (**4**), pp. 1036-1060.
2. Heil, M., Rosler, F., Link, M., and J. Bajric (1998). What is improved if a mental rotation task is repeated – the efficiency of memory access, or the speed of a transformation routine?. In: *Psychological Research*, **61**(**2**), pp. 99-106.
3. Kail, R., & Park, Y.S. (1990). Impact of practice and speed of mental rotation. *Journal of Experimental Child Psychology*, **49**, 227-244.
4. Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, pp. 492-527.
5. Lotz, A., and N. Russwinkel (submitted). *Kognitive Systeme*, Bochum, Germany.
6. Shepard, R.N., and J. Metzler (1971). Mental Rotation of Three-Dimensional Objects. *Science*, **171**, pp. 701-703.
7. Tarr, M. J. & S. Pinker (1989). Mental rotation and orientation-dependence in shape recognition. *Cognition Psycho*logy, **21**, pp. 233-282.
8. Wallace, B. & Hofelich, B. G. (1992). Process generalization and the prediction of performance on mental imagery tasks. *Memory & Cognition*, **20**, 695-704.

Rule Learning from Incremental Presentation of Training Examples: Reanalysis of a Categorization Experiment*

Christina Zeller and Ute Schmid

Cognitive Systems Group, University of Bamberg
An der Weberei 5, 96045 Bamberg, Germany
{Christina.Zeller,Ute.Schmid}@uni-bamberg.de

Categorization is an important cognitive skill that helps us, for example, to make predictions about objects belonging to the same category and to communicate by referring to types of objects with words. There are four main theories that try to explain categorization: the rule-based, the prototype, the exemplar, and the decision-boundary theory (Kruschke, 2008). Furthermore, these approaches can be combined to hybrid theories connecting two or more theories together with criteria when which strategy is applied. In hybrid theories rule-based approaches are often included to explain categorization with logical rules. However, for a long time it was criticized that rule-based theories have shortcomings in explaining typicality effects, that is, that more typical members of a category are treated more efficiently. Recently it was shown that typicality effects can also be explained by a rule-based approach (Lafond, Lacouture, & Cohen, 2009).

Cognitive theories of categorization usually focus on the categorization process and do not make specific assumptions about underlying learning algorithms. Designing such learning algorithms—however, without explicitly focusing on cognitive learning—is the domain of machine learning research (Mitchell, 1997).

In this paper we take a closer look on the incremental process of category learning with a rule-based approach. First, we describe an experiment of human categorization learning which gives evidence that humans use rule-based categorization strategies (Lafond et al., 2009). The authors model rule-based categorization with individual decision-trees (iDT). Afterwards, we present our reanalysis of these data with the focus on incremental learning of iDTs.

Experiment and Decision-Tree Models

In the experiment, participants categorized computer-generated 3D rendered images of lamps in categories A and B (Lamberts, 2000; Lafond et al., 2009). The lamps differed in four binary features (F1–F4): base, upright, shade, and top. For example, the shade can be conical or hemispherical. The resulting 16 stimuli can be represented compactly by a four-digit binary code—e.g., 0000: all features have the attribute value 0; 1010: the attribute value of F1 and F3 is 1. Stimuli

* A big thanks to Daniel Lafond who made the data and the material of the categorization experiment available for us and who helpfully answered our questions.

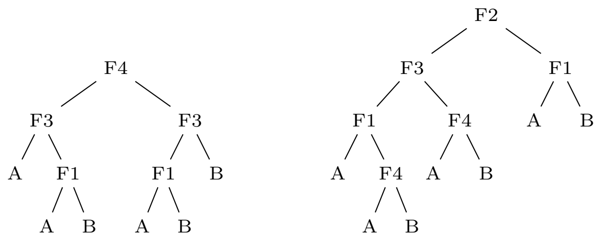


Fig. 1. Decision-trees for Participants 1 (left) and 4 (right). Left branches correspond to attribute values 1, right branches to attribute values 0.

were assigned to the two categories with the Medin and Schaffer 5–4 category structure (cf. Medin & Schaffer, 1978). The mapping of abstract to physical (lamp) features was counterbalanced (i.e. differently) across participants.

The experiment included a training and a process tracing phase. During training, the labeling for stimuli A1–A5 and B1–B4 was learned with trial-by-trial feedback. The stimuli were presented incrementally in blocks of nine stimuli in random order, differing for each participant. Training ended after a participant categorized all nine stimuli in a block correctly three times in a row. In the process-tracing phase the participants performed the *four-questions game* where the participants had to decide which feature should be shown first, second, third and fourth. At each step, the participant could stop the trial by categorizing the stimulus. Participants were instructed not to guess, to use the same strategy they used during training, and to use a minimal number of features.

Data from five participants (P1–P5) where analyzed and results of the four-questions game led to five iDTs to model the categorization. The structure of the iDTs for P1 and P4 are shown in Fig. 1. These iDTs explain the data better than two compared exemplar models (cf. Lafond et al., 2009). However, there are open questions concerning (a) the structure of these iDTs and (b) the process of incremental learning these iDTs. In the following, we will give a closer look to these aspects.

Process of Learning Categories: A Data Reanalysis

When inspecting the iDTs underlying the categorization process of the participants, it shows that the trees have different structures, for example, one has F4 the other F2 as root. Consequently, to develop a model for the incremental learning process, it has to be determined by which criteria individuals select features for inclusion in the iDTs.

For decision-tree (DT) learning algorithms in machine learning a widely used selection criterion is information gain. This measure describes the reduction of entropy in a system when a feature is used to split the data into categories (Quinlan, 1986). However, to calculate information gain the complete set of data and their categories must be known. In contrast, human learning mostly takes place in incremental settings. In the described experiment, after one training block a participant theoretically can calculate the information gain for all four features which would result in: $gain(F1) = gain(F3) = 0.229; gain(F2) = 0.007; gain(F4) = 0.091$. Features F1 and F3 have the highest information gain. However, only the iDT of P5 has F3 as root, whereas in the iDTs of P1, P2, and P3 feature F4 is root and of P4 it is feature F2. Nevertheless, for all further nodes of all iDTs always a feature with the highest information gain for all stimuli in this path of the tree was used.

An explanation for not using the feature with the highest information gain as root could lay in the presentation of the material in the first block in the training phase. If a participant only has seen the stimuli A4 (1101), A5 (0111) and B1 (1100) feature F4 discriminates the stimuli correctly in Category A and B. Based on this observation we defined an information gain related measure '*igain*' that only includes the stimuli that are known by the participant so far. Since this measure only takes into account a limited number of stimuli, the relative frequencies do not have high validity. Therefore, while the relative frequencies in the entropy measure are interpreted as probabilities, this interpretation is not justified for *igain*. For a set of stimuli at a given time S_t and a category c, the relative frequency is calculated as $r(S_t, c) = \frac{|S_{t,c}|}{|S_t|}$ where $S_{t,c}$ is the set of stimuli belonging to category c. Entropy is calculated in the usual way with the relative frequencies instead of probabilities $H(S_t) = -r(S_t, A) \times log_2 r(S_t, A) - r(S_t, B) \times log_2 r(S_t, B)$. The *igain* measure is defined as

$$igain(S, F, t) = H(S_t) - \frac{|S_{t,F^0}|}{|S_t|} \times H(S_{t,F^0}) - \frac{|S_{t,F^1}|}{|S_t|} \times H(S_{t,F^1})$$

with S_{F^0} as stimuli with attribute value 0 for feature F and S_{F^1} as stimuli with attribute value 1.

In Table 1 the analysis of the data with respect to *igain* is shown. The first non zero values occur when stimuli of both category A and B have been seen. Therefore, P2 needed a longer sequence to get a non-zero *igain* value. For the iDTs of P1–P4 always a feature with the highest *igain* was used. For P5 the only exception is the feature used in the root. For the stimuli sequences used in the experiment, there are some cases where two or more features have the same *igain*, that is, we have no unique criterion for selecting a specific feature. Since it cannot be supposed that participant have a perfect memory, an additional selection criterion could be the visual salience of a feature.

Identifying principles which explain the order of features in iDTs addresses one aspect of the learning process. Additionally, it is an open question whether human learners follow a strictly incremental strategy for constructing and refining rules. We compared the number of trials, that is the number of learning steps, participants needed until all stimuli are correctly classified with the number of steps necessary for the DT algorithm CAL 2 (Unger & Wysotzki, 1981) when features were presented as evident in the iDTs. Results show that CAL 2 does need fewer steps than the participants (see Table 1). Obviously four of the five participants do not follow a purely incremental strategy as used by CAL 2.

Table 1. Incremental information gain and learning steps (explanation see text).

Participant	Incremental information gain					Learning steps		SC
	Sequence	F1	F2	F3	F4	CAL 2	Participant	
P1	B3, A2	1	0	1	1	25	77	0.680
P2	A3, A5, A4, B2	0.311	0.123	0.123	0.811	23	63	0.696
P3	B2, A5	0	0	0	1	17	161	0.882
P4	A5, B3	0	1	1	0	24	651	0.708
P5	B2, A3	1	1	0	1	28	101	0.607

One possible alternative model is that participants completely reject partially learned rules and start with a new hypothesis. This was also proposed by Unger & Wysotzki (1981) who assume a meta-strategy by which humans first focus on simple rules and switch afterwards to conjunctions or disjunctions of features. Furthermore, if human learners followed the incremental strategy, the classification decisions during learning should correspond between participants and the learning algorithm. We assessed sequence consistency (SC) by calculating the proportion of matches between participants' response sequences and CAL 2 until CAL 2 had learned the complete iDTs (see Table 1). Results show only moderately high correspondences with the highest correspondence for P3.

Our reanalysis shows that—although categorization behavior of the participants can be reconstructed by iDTs—the process of learning cannot be plausibly modeled by the incremental DT algorithm CAL 2 for four of the five participants. In a next step, we plan to conduct an experiment using the same stimuli but different learning set-ups to gain more insight in rule-based classification learning. There we will explore selection criteria for features taking into account the proposed *igain* measure and visual salience (cf. utility values proposed by Lamberts 2000). Furthermore, the use of different meta-strategies will be researched by assessing participants' categorization rules after each learning step.

References

Kruschke, J. K. (2008). Models of categorization. In R. Sun (Ed.), *Computational psychology* (pp. 267–301). Cambridge University Press.

Lafond, D., Lacouture, Y., & Cohen, A. L. (2009). Decision-tree models of categorization response times, choice proportions, and typicality judgments. *Psychological Review*, *116*, 833–855.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Mitchell, T. M. (1997). *Machine learning.* New York, NY: McGraw-Hill.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106.

Unger, S., & Wysotzki, F. (1981). *Lernfähige Klassifizierungssyssteme (Classification Systems Being Able to Learn).* Berlin, Germany: Akademie-Verlag.

Multisensory Conflict yields Adaptation in Peripersonal and Extrapersonal Space

Johannes Lohmann and Martin V. Butz

Cognitive Modeling, Department of Computer Science, Faculty of Science, University of Tübingen, Germany
{johannes.lohmann, martin.butz}@uni-tuebingen.de

Abstract. Spatial representations are acquired through active interaction with the environment and are based on a multisensory integration mechanism that combines visual, tactile and proprioceptive information. The weighting of different modalities depends on their reliability and changes from peripersonal to extrapersonal space. In a virtual reality setup we investigated whether conflict between visual and proprioceptive information regarding the hand position yields adaptation of spatial representations. Our results show a stronger bias towards the manipulated visual information for localizations in extrapersonal space. The data is consistent with the assumption that peripersonal space is more strongly grounded in proprioceptive than visual information, compared to extra-personal space.

Keywords: Spatial Perception, Peripersonal Space, Virtual Reality

1 Introduction

Active interaction with the environment shapes the way we perceive the space around us and internal models used to predict action outcomes originate from these interactions [1]. Each motor command provides a variety of visual, tactile, proprioceptive and acoustic sensations, which are integrated into a coherent percept by means of a maximum likelihood integrator [2]. Especially for the immediate space around the body - the so-called peripersonal space - the close relation between motor codes, vision and proprioception has been shown [5]. Hence, the internal representation of peripersonal space is not defined in terms of a Cartesian metric, but in terms of sensorimotor functionality. With increasing distance from the body, the weighting of visual and proprioceptive information in the spatial representation changes. Longo & Lourenco [4] could show that the representation of extrapersonal space is dominated by visual information. This transition in the weighting of visual and proprioceptive information is continuous and scales with arm-length.

To investigate the weighting of information and how the spatial representations are formed during sensorimotor interactions, an active manipulation of the mapping between modalities is required. Classic methods to introduce multisensory conflict - like the rubber hand illusion - require participants to remain motionless. Virtual reality (VR) setups offer a possible solution. We manipulated the mapping between visual and proprioceptive hand position to investigate whether the integration of the conflicting sensory information yields adaptation of spatial representations. Participants had to perform a bimanual task during which the visual hand representation was shifted, resulting in a correction of the actual hands, to maintain the target position in the VR. The mismatch between proprioception and vision should yield an adaptation in the representations of peri- and extrapersonal space, which we measured via localization tasks in near and far space. We expected stronger mislocalizations for the far space since it should be adapted according to the manipulated visual impression. To further explore the role of visual saliency, we hid the virtual hand models during the localization in half of the trials.

2 Method

Participants. 33 students from the University of Tübingen participated in the study (22 males). Their age ranged from 18 to 30 years (M = 21.7, SD = 2.5). Participants were told a cover story to keep them naive to the purpose of the study. After the experiment, participants were debriefed and offered the opportunity to withdraw their data.

Virtual Reality Setup. Participants were equipped with an Oculus Rift © DK2 stereoscopic head-mounted display. Hand motions were captured with a Leap Motion © near-infrared sensor, placed 30cm in front of the participants on a table. The VR scenario put participants in a static mountain scenery, with a basket at the outer right corner of their reachable task space. During the experiment a flower spawned at the center of the scene and participants had to pick the petals and put them into the basket (see Fig. 1, panel A).

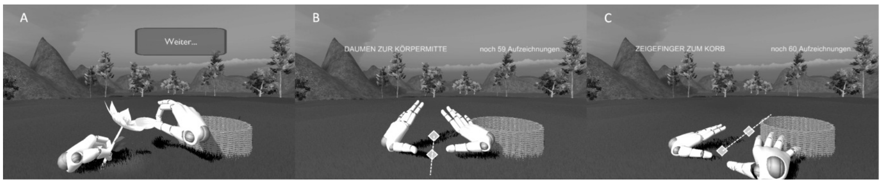


Fig. 1. Panel A: Object interaction task. Panel B: Self-Localization, diamonds indicate palm and thumb centroids, respectively. Panel C: External Localization, diamonds indicate palm and index finger centroids, respectively.

Procedure. In each trial, participants had to perform three tasks. First a localization, second the object interaction during which the visual offset was applied to the hand model, finally they repeated the localization task. The experiment consisted of two blocks, each consisting of 12 trials.

Localization. Participants had to locate themselves and an external reference within the scene by pointing to the reference with both hands. For the self-localization, participants were instructed to point with the tip of their thumbs to themselves. In case of the external reference, participants were instructed to point at the basket with their index fingers. The two types of localization are displayed in Fig. 1 (panel B and C). The experiment was divided into two blocks. In one block, the hand model was displayed during the localization, while it was hidden in the other block.

Object Interaction. After the initial localization was accomplished, a flower bloomed in the center of the scene. Participants were instructed to pick as many petals as possible and to put them into the basket. In order to so, they had to grab the stem with the left hand and to pick the petals with the right hand (see Fig. 1, panel A). During task the offset between visual and felt hand position was introduced. The offset was introduced gradually and only while the hands were moving. Participants complied with the task in general, collecting 4.5 petals on average per trial (SD = 1.4).

Design. To test systematic effects on the localization performance, we used a 2 $\times$ 2 design with the factors visibility (hand visible during localization or not) and reference (pointing towards external reference or towards self). We derived three dependent measures for the quantitative analysis. The *palm drift* refers to the difference between the hand centroid in the pre- and post-localization and indicates adaptation of the spatial representations of the hands. A shift in hand position does not necessarily lead to mislocalization, the *angular disparity* quantifies the adaptation of the hand rotation from the the pre- to the post-localization which compensate possible drifts. To assess changes in the actual localization, the *positional discrepancy* is the difference between the positional estimates in the pre- and post-localization.

3 Results

Data were analyzed with 2 (hand visibility) $\times$ 2 (positional reference) repeated measure ANOVAs. Results are displayed in Fig. 2. For all variables, main effects for hand visibility and reference were obtained, the respective interaction was only significant for the *angular disparity*. For all conditions and measures, means differed significantly from zero, the only exception being the positional discrepancy in case of invisible hands and self-localization.

4 Discussion

We dissociated visual and the proprioceptive hand position in a VR setup and tested whether the introduced dissociation affected localization performance. To manipulate the saliency of visual and proprioceptive information, we let the participants perform the localization task either with visible, or invisible virtual hands. The data implies that participants stuck to the shifted center of their

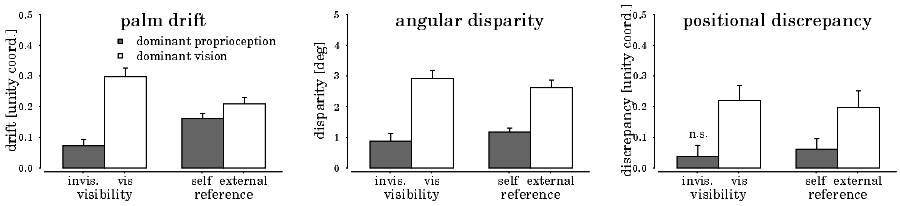


Fig. 2. Main effects for hand visibility (left) and localization reference (right). Both main effects are significant for all measures. Bars with gray background indicate conditions where the localization relied more on proprioceptive information. All means differed significantly from zero, except in case of positional discrepancy and invisible hands (this condition is marked with "n.s."). Please note that the scale for angular disparity indicates angles in degrees, while for the two other measures, the y-axis represents units in Unitys' coordinate system.

hands, but partially compensated this shift by an according rotation of their palms in the localization tasks. Results with respect to the positional discrepancy show how the participants adapted their location estimate in a systematic way, reflecting the introduced visual offset. The only exception was the combination of invisible hands and self-localization - the most proprioceptive condition so to say. Our results show how multisensory conflict yields adaptation of the spatial representation of far space and, to a smaller degree to an adaption of the self-localization. The results dovetail with earlier work that showed a different weighting of proprioceptive and visual information in peripersonal and extrapersonal space [3]. Furthermore, the results highlight the dynamic nature of spatial representations. Earlier studies have shown the fast remapping of peripersonal space in case of tool-use, our results extend these findings by showing the remapping of ego- and allocentric frames of reference due to sensorimotor interaction.

Active manipulation of spatial representations in VR allows to study aftereffects on spatial reasoning and spatial compatibility effects. This will provide an even deeper understanding how spatial representations are rooted in the sensorimotor system and how they affect higher cognitive functions.

References

1. Barsalou, L.W.: Grounded cognition. Annual Review of Psychology 59, 617–645 (2008)
2. Ernst, M.O., Bülthoff, H.H.: Merging the senses into a robust percept. Trends in Cognitive Sciences 8(4), 162–169 (2004)
3. Longo, M.R., Lourenco, S.F.: On the nature of near space: Effects of tool use and the transition to far space. Neuropsychologia 44(6), 977–981 (2006)
4. Longo, M.R., Lourenco, S.F.: Space perception and body morphology: extent of near space scales with arm length. Experimental Brain Research 177(2), 285–290 (2007)
5. Macaluso, E., Maravita, A.: The representation of space near the body through touch and vision. Neuropsychologia 48(3), 782–795 (2010)

Matching Matchboxes: Co-actors Create Non-conventional Communication Systems for Joint Action

Cordula Vesper, Tiffany Morisseau, Günther Knoblich, Dan Sperber

Department of Cognitive Science, Central European University, Budapest, Hungary
{vesperc,morisseaut,knoblichg,sperberd}@ceu.edu

Abstract. We investigated how non-conventional communication systems are created in joint action. Results from two experiments, in which a knowledgeable co-actor informed a partner about object categories, suggest that marking object categories ostensively may be crucial for establishing successful communication.

1 Background

When multiple people coordinate their actions towards performing a common goal [1], they often rely on different forms of communication to facilitate their joint action. Spoken language is an obvious candidate that often eases social interaction [2]; but people also rely on non-verbal communication such as gestures [3]. While the functionality of such forms of communication for coordination is undebatable, less is known about how people explore new ways of providing information to each other in cases when conventional communication systems such as formal language are unavailable [4].

The research field of experimental semiotics provides a tool to study the emergence of new communication systems in a controlled lab environment [5]. Typically, participants are placed in a situation where they cannot rely on conventional forms of communication and instead need to use other ways of interacting. This research shows that people are generally good in creating new and stable communication systems; however, the success rate varies strongly between individuals.

In many cases of joint action, information relevant for performing a task together is perceptually retrievable and it may be sufficient for co-actors to highlight or enhance it. In other cases, the relevant information can be indicated by means of an unambiguous conventional signal or iconic gesture. But what happens when a crucial piece of information can neither be displayed nor unambiguously signaled? We predicted that, in such a case, participants would resort to *ostensive communication* [6,7], i.e. that they would give evidence of their intention to communicate relevant information in a way that would allow the addressee to infer what information they intended to communicate.

2 The present study

The present study aimed at exploring under which conditions and in which ways ostensive communication may emerge in new communication systems for joint action. To that end, we designed a task (Fig. 1) in which objects belonging to different categories had to be matched between an informed communicator ('Leader') and an uninformed receiver ('Follower'). The objects (plain matchboxes) were distinguishable on different levels: Whereas a 'colored' matchbox could be distinguished from a 'normal' matchbox based on overtly perceptible features, the distinguishing features of a 'special' box were hidden and therefore not directly accessible for an observer. The Leader's task in the present study was to inform the Follower about these object categories so that both co-actors could pick and place the same object type into a common target area. The Follower's task was to understand and use the communicative cues provided by the Leader.

Fig. 1. Experimental setup. The Follower's stack is visible on the left, whereas the Leader's stack is behind the separation on the right side. The central target area is visible to both partners.

The challenges in this joint matching task were that, first, no conventional form of communication (i.e. speech) was allowed, that, second, co-actors only shared a minimal part of the environment (i.e. they could only see the common target area and the partner's hand movements while the partner's upper body and face were hidden from view) and that, third, no direct feedback about communication success was provided. How would co-actors solve this coordination problem?

We hypothesized that Leaders would invent a new communication system by using their movements during or after placing a matchbox in order to inform Followers to what category the box belongs. Specifically, we predicted that

1. 'colored' boxes would either be simply placed or the distinguishing perceptual feature would be enhanced to support its detection by Followers,
2. 'normal' boxes would not be marked by a specific gesture, and
3. 'special' boxes would be indicated by means of a form of ostensive communication.

Our main interest was to understand how Leaders would distinguish 'special' from 'normal' boxes and how that would be reflected in Followers' success to understand the intended meaning and choose the same matchbox type. We designed the task in a way that what made boxes 'special' (i.e. a hidden rectangle drawn inside) could not be directly shown or unambiguously indicated by means of an iconic gesture (such as tracing a rectangle in the air since all the surfaces of all the boxes were themselves rectangular) – unless the gesture was taken to be an act of ostensive communication. We predicted that pairs resorting to ostensive communication to indicate specialness (with or without some degree of iconicity) would be more successful in establishing a communication system than pairs not making use of ostensive communication.

3 Results

In Experiment 1, in which we analyzed video data and self-reports of 24 participants in randomly assigned pairs, 9 out of 12 pairs successfully established a stable communication system (i.e. they had less than 20 % category mismatches). In all but one of those successful pairs, Leaders did not mark the ‘colored’ or the ‘normal’ boxes but invented an ostensive gesture indicating an intention to communicate for the ‘special’ matchbox category (6 Leaders used purely ostensive cues such as tapping on the box; 3 Leaders used an ambiguous iconic gesture such as drawing a virtual rectangle; 1 Leader used a combination of both). Followers understood the partner’s communicative intention independent of what type of gesture the Leader used. Thus, communication could be successful even if the Leader’s precise intention in using a partially iconic gesture and the Follower’s interpretation of it differed. For example, in one pair, an intended iconic box opening gesture (a pinching movement of thumb and index finger) was not understood iconically but was nevertheless interpreted as signaling the ‘special’ box.

Given this result of Experiment 1, where solely communicating ‘special’ was sufficient to establish a clear contrast to ‘colored’ and ‘normal’, we performed a second experiment with 24 new participants that investigated whether increasing the number of required gestures would influence the Leader’s communication and the Follower’s understanding of it. To this end, we changed the perceptual feature of the ‘colored’ box such that it could only be detected if the Leader actively showed it to the Follower. Specifically, whereas in Experiment 1, the distinguishing feature, a red line, was placed on the side of the matchbox, in Experiment 2 it was underneath the box such that it was hidden from view in the target area. In all other respects, the experiments were identical.

Experiment 2 revealed a successful communication system in 7 out of 12 pairs. Most of the Leaders (5 out of 7) informed Followers about the ‘colored’ matchbox by simply showing them the bottom of the box before placing it. (The 2 others invented a separate iconic gesture.) ‘Special’ was again marked by a gesture that either had an iconic component or was purely ostensive. Interestingly, the key difference to the first experiment was how Leaders treated the ‘normal’ matchbox category. Now only 2 of the 7 Leaders chose not to use any specific gesture for it. The majority created a communicative cue aimed at enhancing the critical difference between ‘normal’ and ‘special’, i.e. the absence of an object feature in ‘normal’. Given that this was not in principle needed to distinguish the different matchbox types, it shows that increasing the number of required gestures influenced whether Leaders relied on treating ‘normal’ as a default category not requiring a separate gesture. In other words, our between-experiment manipulation biased Leaders in Experiment 2 to code a full communication system with three separate gestures. The ‘special’ category, finally, was again signaled with a gesture that was either iconic (1 Leader) or purely ostensive (5 Leaders), although one Leader marked it with a showing gesture similar to 'colored', thereby demonstrating the absence of the distinguishing color feature (there was no gesture for 'normal' in this case).

Of central interest was also why a number of pairs failed to establish a communication system and consequently did not succeed in matching all objects. Of the 8 pairs in Experiments 1 and 2, three failed because Leaders provided unclear signals and two more because Followers did either not understand the Leaders' intention or did not take it into account for their own behavior. The remaining three cases are most interesting because they demonstrate incidents of misunderstanding. These pairs consistently mixed up the 'normal' and the 'special' categories without realizing it. In two cases, the Leader's iconic gesture (stroking the inside hand with the thumb to symbolize the hidden feature) or ostensive cue (shaking the box) for the 'special' category was interpreted as a gesture of emptiness ('normal'). In a third case, the Leader chose to signal feature absence for the 'normal' category and did not mark 'special' at all which the partner misunderstood because of an expectation that specialness would be marked. Especially this latter case suggests that providing cues that clearly indicate a Leader's *communicative intention* are important for coordination success – possibly more than the specific form of the gesture.

4 Conclusion

The present study investigated how partners in a joint action create and rely on non-conventional communication systems to convey information about object categories. Our findings demonstrate that marking object categories ostensively may be beneficial or even crucial for establishing a successful communication system. Moreover, a need to enhance a perceptual object feature for a joint action partner influences the overall number of specific codes used in the communication system. Taken together, the present work suggests that for people performing a joint action, ostensive communication may provide a powerful mechanism to achieve real-time action coordination when task-relevant information can neither be perceptually highlighted not unambiguously communicated by means of conventional or iconic signals.

5 References

1. C. Vesper, S. Butterfill, G. Knoblich, G., and N. Sebanz, "A Minimal Architecture for Joint Action", *Neural Netw*, vol. 23, pp. 998-1003, 2010.
2. H. H. Clark, "Using Language", Cambridge, England: Cambridge University Press, 1996.
3. A. Özyürek, S. Kita, S. E. Allen, R. Furman, and A. Brown, "How does Linguistic Framing of Events Influence Co-speech Gestures? Insights from Crosslinguistic Variations and Similarities", *Gesture,* vol. 5, pp. 219-240, 2005.
4. T. C. Scott-Phillips, "Speaking Our Minds", Palgrave Macmillan, 2014.
5. B. Galantucci, "Experimental Semiotics: A New Approach for Studying Communication as a Form of Joint Action", *Topics Cog Sci,* vol. 1, pp. 393-410, 2009.
6. D. Sperber and D. Wilson, "Relevance: Communication and Cognition", Oxford/Cambridge: Blackwell Publishers, 1995.
7. G. Csibra, and G. Gergely, "Natural Pedagogy", *Trends Cog Sci,* vol. 13, pp. 148-153, 2009.

Rethinking the Keystroke-Level Model from an Embodied Cognition Perspective

Marc Halbrügge

Quality & Usability Lab, Technische Universität Berlin
Ernst-Reuter-Platz 7, 10587 Berlin
marc.halbruegge@tu-berlin.de

Abstract. Since its first presentation in the 1980s, the Keystroke-Level Model (KLM) has been successfully applied in the domain of human-computer interaction. Although being much simpler than the GOMS technique, it still allows sufficiently accurate predictions of task completion times for a given user interface (UI) and task. A closer look at the data however reveals the necessity of adaption to UI paradigms that did not exists when Card et al. formulated the KLM. Graphical UIs allow the use of embodied strategies which lead to extensions to the original KLM heuristics. The resulting model is evaluated based on a reanalysis of data from three different studies.

Keywords: Human-Computer Interaction; Keystroke-Level Model; Cognitive User Model

1 Introduction

More than 30 years have passed since the publication of the seminal "Psychology of Human-Computer Interaction" [1]. Card, Moran, and Newell's approach of using a computer metaphor to describe the characteristics of human computer users (i.e., the Model Human Processor; MHP) proved very successful in the following. By assigning computation speeds to three perception, cognition, and motor processors that work together, the MHP is capable of explaining many aspects of human experience and behavior.

Card et al. derived the GOMS technique (Goals, Operators, Methods, and Selection rules) from the MHP, which provides fine grained predictions of task completion times but is seldom applied because it is rather hard to learn [6]. An easier solution is provided by a simplified version of GOMS, the Keystroke-Level Model (KLM). The KLM mainly predicts task completion times by dividing the necessary work into physical and mental actions. The physical time (e.g., a mouse click) is predicted based on results from the psychological literature and the mental time is modeled using a generic "Think"-operator M that represents each decision point within an action sequence. M takes about 1.35 s, which has been determined empirically by Card and colleagues. While the generic M operator may oversimplify human cognition, predictions based on the KLM are easy to obtain using computer tools and are also sufficiently accurate (i.e., within 10% absolute percent error; [7]).

1.1 Display-Based Difference-Reduction

The main paradigm studied by Card et al. was document editing and professional secretary work using command-based text editors. Such editors rely heavily on modes (e.g., users cannot add text outside some 'insert' mode) and are operated based on memorized command strings (examples for Vim[1]: `":q!"` or `"10dd"`). Users that behave according to the KLM are assumed to have perfect knowledge in-the-head [8]. If the visual presentation of the interface does not keep up with their actions, they do not wait for the UI, but type-ahead the next command.

Since then, the software landscape has changed dramatically. Direct manipulation is the prominent paradigm since the introduction of graphical UIs (GUI). One major effect of these changes is that today's systems aim at providing as much knowledge in-the-world [8] as possible. Thereby, the cognitive demand of a task can be significantly lowered. The actual work is shifted from the cognitive to the perceptual domain. As Gray [2] has pointed out, the presentation of the current state of a system in its GUI reduces the necessity of place-keeping in the users' memory. He has called the resulting user strategy *display-based difference-reduction* (DBDR). When using DBDR, users do not follow a memorized action sequence, but choose their actions based on whether and where the GUI visually differs from an intended target state.

2 Extended KLM Heuristics

The DBDR strategy follows an embodied perspective by proposing that humans use the information provided on a GUI to minimize cognitive load [11]. Applied to the KLM, we can derive two assumptions about user behavior that could not have been observed at Card et al.'s time.

- Periods of blank screen block the users. The original KLM assumes that blank screen time can be used for memory retrievals. This does no longer apply when users follow the vision-based DBDR strategy instead.
- Especially when interacting with slightly inaccurate systems, e.g., touch devices, users have to visually monitor the GUI to check whether their actions have had the intended effects.

3 Empirical Validation

The KLM extensions presented above were originally developed and validated based on two independent experiments using a virtual kitchen assistant [9]. The paradigm consisted of a set of simple to fairly complex tasks that the participants had to complete using that assistant, e.g., "Search for German main dishes and select lamb chops" or "Create a shopping list for five servings". We subsequently used the same paradigm in several experiments within a research project about human error [3–5]. Those experiments differ from [9] in several respects, e.g., the devices used [3], additional data recorded [4], and the presence of additional tasks [5]. None of these experiments has

[1] http://www.vim.org/

been analyzed with regard to task completion time before. They should therefore provide a good opportunity to test the generalizability of the KLM extensions.

			original CogTool			extended KLM		
Experiment	**Date**	**N**	R^2	RMSE	MLSD	R^2	RMSE	MLSD
Pretest [9]	2013-05	10	.597	0.39 s	4.6	.995	0.13 s	1.4
Original Validation [9]	2013-11	12	.425	0.61 s	20.0	.927	0.47 s	13.8
Error Classification [3]	2014-07	20	.485	0.43 s	10.8	.930	0.25 s	10.2
Eye-Tracking [4]	2015-01	24	.440	0.51 s	16.3	.904	0.32 s	15.7
Multi-Tasking [5]	2015-05	12	.595	0.43 s	6.0	.975	0.21 s	5.6

Table 1. Goodness-of-fit of the original CogTool model and after application of the extended KLM heuristics proposed in this paper.

The validity of the extended heuristics is examined by comparing the goodness-of-fit of classical KLM predictions (using CogTool; [7]) with the ones obtained after applying the new heuristics (see Table 1 and Figure 1). Four different types of user clicks are compared that feature the effects of the extended heuristics. The *new screen* condition differs from the *other group* (of elements) condition by a blank period during the transition to another GUI screen. The *same group* condition accordingly denotes clicks within a set of semantically grouped elements, while *same button* refers to subsequent clicks on the same element. The predictions for *same group* contain monitoring time that is masked by think time in the *other group* and *new screen* conditions (see also [9]).

Besides R^2 and RMSE, we are giving the Maximum Likely Scaled Difference (MLSD; [10]) which scales the deviation of a model's prediction from the empirical mean using the length of the confidence interval, thereby providing fairer comparisons of different models.

4 Discussion and Conclusions

The good fit (R^2 always above .9) of the extended model provides evidence that the embodied DBDR strategy is actually used by today's computer users. This change should be reflected by extending the KLM heuristics accordingly.

Acknowledgments This work was supported by DFG grant MO 1038/18-1 ("Automatische Usability-Evaluierung modellbasierter Interaktionssysteme für Ambient Assisted Living").

References

1. Card, S.K., Moran, T.P., Newell, A.: The Psychology of Human-Computer Interaction. Erlbaum Associates, Hillsdale, New Jersey (1983)

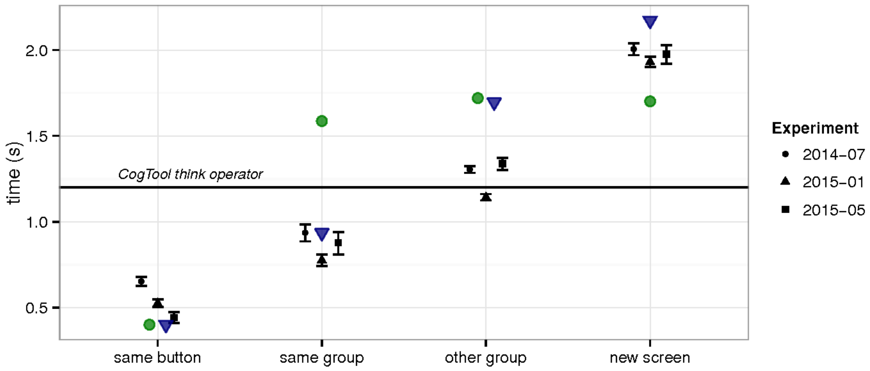


Fig. 1. Time per click with model predictions. See text for detailed explanations of the four click types. ● denote original CogTool predictions, ▼ denote extended KLM predictions, error bars denote 20% trimmed means with bootstrapped 95% confidence intervals.

2. Gray, W.D.: The nature and processing of errors in interactive behavior. Cognitive Science 24(2), 205–248 (2000)
3. Halbrügge, M., Quade, M., Engelbrecht, K.P.: A predictive model of human error based on user interface development models and a cognitive architecture. In: Taatgen, N.A., van Vugt, M.K., Borst, J.P., Mehlhorn, K. (eds.) Proceedings of the 13th International Conference on Cognitive Modeling. pp. 238–243. University of Groningen, Groningen, the Netherlands (2015)
4. Halbrügge, M., Quade, M., Engelbrecht, K.P.: Cognitive strategies in HCI and their implications on user error. In: Proceedings of the 38th Annual Meeting of the Cognitive Science Society (2016), accepted
5. Halbrügge, M., Russwinkel, N.: The sum of two models: How a composite model explains unexpected user behavior in a dual-task scenario. In: Proc. ICCM 2016 (2016), accepted
6. John, B.E., Jastrzembski, T.S.: Exploration of costs and benefits of predictive human performance modeling for design. In: Proceedings of the 10th International Conference on Cognitive Modeling. pp. 115–120. Philadelphia, PA (2010)
7. John, B.E., Prevas, K., Salvucci, D.D., Koedinger, K.: Predictive human performance modeling made easy. In: CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 455–462. ACM Press, New York, USA (2004)
8. Norman, D.A.: The psychology of everyday things. Basic books, New York, NY (1988)
9. Quade, M., Halbrügge, M., Engelbrecht, K.P., Albayrak, S., Möller, S.: Predicting task execution times by deriving enhanced cognitive models from user interface development models. In: Proceedings of the 2014 ACM SIGCHI Symposium on Engineering Interactive Computing Systems. pp. 139–148. EICS '14, ACM, New York, NY, USA (2014)
10. Stewart, T.C., West, R.L.: Testing for equivalence: a methodology for computational cognitive modelling. Journal of Artificial General Intelligence 2(2), 69–87 (2010)
11. Wilson, M.: Six views of embodied cognition. Psychonomic bulletin & review 9(4), 625–636 (2002)

Communicative signaling and self-other distinction: Next steps for an embodied hierarchical model of dynamic social behavior and cognition

Sebastian Kahl and Stefan Kopp

Social Cognitive Systems Group, Faculty of Technology,
Bielefeld University, Inspiration 1, 33619 Bielefeld, Germany
{skahl,skopp}@uni-bielefeld.de

1 Introduction

In order to engage in a continuous social interaction, participants must be able to dynamically understand, predict, and influence the mental states and actions, so as to enable a process of efficient and interactive grounding of shared meaning. We follow the argument that the mentalizing network and the mirror-neuron system in our social brain together provide the basis for these abilities [1, 2]. However, how these systems exactly operate and how they work together is still unclear. Building on previous work on the interplay of mentalizing and mirroring in embodied communication, here we lay out next steps towards an embodied hierarchical model of dynamic social behavior and cognition. The proposed next steps target the early and reliable self-other distinction and integration in the sensorimotor system, which in turn informs mentalizing, so that it can distinguish between own and other's beliefs in complex situations of simultaneous action perception and production. Also, we propose that social cognitive systems informed in such a way have the information to allow for strategic signaling behavior by selecting actions necessary to make their action goals easier to disambiguate, and thus to communicate efficiently and successfully.

2 Current model

In a first step towards this goal, we developed a model of two distinct networks of the human social brain - mentalizing and mirroring - which allows them to interact during embodied communication. The model connects a mentalizing system based on simple heuristics for attributing and inferring different orders of belief about own and other's mental states, with a hierarchical predictive processing model of online action perception and production based on the common coding of underlying action representations [3]. To investigate the role of mentalizing and mirroring interacting in inter-agent coordination and to test the model, we conducted simulation experiments in which two virtual agents were each equipped with this model. Different mentalizing capacity configurations were tested, as well as different noise conditions, thus influencing the robustness of the communication. The agents engage in non-verbal communication behavior to which the embodied action representations in the mirroring system can resonate because of their close coupling of perception and production, while taking uncertainty from noise into account. Resonating action representations inform the mentalizing system, which in turn can guide successful interaction. Results from our simulations on this first model demonstrate how mentalizing can afford higher robustness of communication by enabling interactive grounding processes.

3 Next steps

Although our model was able to act upon and infer beliefs about own and other's mental states, it could only produce or perceive an action at a time. Of course, this is a special case of interaction that can occur, but in our dynamic world our social brain has to cope with simultaneous interaction with multiple partners, as well as simultaneous production and perception of actions. As a starting point for an account of simultaneous action and perception, the first step is to enable early self-other distinction within the sensorimotor system. Being able to run predictive sensorimotor processes for both self and other selectively provides the basis for the next step: enabling the mentalizing and mirroring system to plan social actions towards achieving our communicative goal. This allows for communicative signaling, in which a motor act (signal) is being strategically adjusted in order to maximize the expected probability of successful reception.

As research into schizophrenia has shown, reliable and early self-other integration and distinction is important not only for the correct attribution of a sense of agency, but also in turn for the correct attribution of intentions and emotions in social interactions [4]. Two major mechanistic models of self-other integration and distinction have been identified. One model, which makes use of people's ability to precisely predict the sensory consequences of their own actions, allows to decrease the intensity of incoming signals by "sensory attenuation", which enables people to distinguish between self-caused actions and their outcomes and those actions and outcomes caused by others. Research even suggests that sensory attenuation correlates with activation in the mirror neuron system [4], and that such attenuation increases during interaction with other people [5]. The second mechanism, which is also influenced by the prediction of action-outcomes, allows for the integration of sensory signals from multiple modalities during a "temporal binding window" [6], which selects perceived actions and their outcomes for integration as long as they occur within a narrow temporal window. Because we have more experience in predicting our own body that window is more narrow for own action-outcomes, than for other people's actions. Being able to make such a distinction allows people to monitor, infer and distinguish between causal relations for own and other's behavior.

How can these mechanisms work in unison to allow us to make such distinctions, even in situations where we simultaneously observe another person perform an action while producing one ourselves?

It is now widely agreed upon, that actions share a common representation for production and perception. Of course, such a common representation is ecological for the brain, since having to sustain multiple representations for one action were quite costly, but also such a common coding can lead to problems like interferences. Such interferences were observed when in a simultaneous action perception and production scenario where the perceived action would be incongruous to the produced one, led to measurable interferences, as a slight mix of the observed action with the produced one [7]. The previously mentioned sensory attenuation can shed light on this effect, since it attenuates predicted action-outcomes only to the degree that we trust the prediction of an action. From a predictive processing perspective, the simultaneous incongruent perception and production of actions would probably lead to strong prediction error signals [8], but a mechanism of sensory attenuation for self-caused predicted action-outcomes can minimize the influence of such a prediction error. However, the observable interferences still indicate that activations due to production or perception can influence each other. Thus, by means of attenuation the sensorimotor system is able to produce and get feedback for own actions while simultaneously perceive actions of others.

This ability to distinguish between own and other's behavior can inform processes that infer beliefs about other people, i.e. a communicative intention that I want an interlocutor to understand, and her behavior that gives me a clue about her understanding of this. This inferred mental state of others, together with my own mental state and a communicative goal, are information that can be used to make following actions and their underlying communicative intent maximally distinguishable from other possibly plausible interpretations, or to compensate for noise. This can be done by communicative signaling as an attempt to strategically alter one's own action kinematics to better achieve the communicative goal [9]. The concept of communicative (or strategic) signaling entails the question on how such alterations are constructed. We will model signaling on an exemplar based approach, where from a set of available actions a selection is made to produce the most distinguishable.

These mentioned mechanisms for self-other distinction, and the resulting ability to make communicative signaling alterations to own behavior are the next steps to be implemented in our embodied hierarchical model of social behavior.

To then be able to achieve self-other distinction in our embodied motor system and make use of the collected information, we plan to test and account for four scenarios that involve embodied agents interacting in a simulated environment, and that require successively increasing abilities for self-other distinction and signaling: In scenario 1, the system will just produce and perceive its own action, to test whether attenuation and temporal binding are working properly. Scenario 2 will test the system's ability to basically face itself in a virtual mirror, so that its own action-outcome and an identical, reflected action-outcome will need to be distinguished. In scenario 3, the system will face a second system, producing a similar, but not equal action, to test its ability to distinguishing between action-outcomes of self and other, as well as to trigger communicative signaling. As in scenario 4, a similar setup will be used, but the second system will produce a completely different action, to test the system's ability to distinguish both actions and trigger communicative signaling as well.

4 Outlook

We are confident that the next steps we have laid out here will propel research towards an embodied hierarchical model of social behavior. We expect this model to provide novel and detailed accounts for several predicted phenomena: First, well known actions are less prone to interference from simultaneously perceived actions. This is due to the increased attenuation for action-outcomes of well known actions and the resulting decreased influence of prediction errors. Second, an exemplar based approach to signaling will allow for a wide variety of possible strategic signaling, which is only limited to the number of actions learned and experienced by the system.

Acknowledgements This research/work was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

References

1. Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., Vogeley, K.: Toward a second-person neuroscience. Behav Brain Sci. 36, 393414 (2013).
2. Sperduti, M., Guionnet, S., Fossati, P., Nadel, J.: Mirror Neuron System and Mentalizing System connect during online social interaction. Cognitive Processing. 15, 307316 (2014).
3. Kahl, S., Kopp, S.: Towards a Model of the Interplay of Mentalizing and Mirroring in Embodied Communication. In: Airenti, G., Bara, B.G., and Sandini, G. (eds.) Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science. 1419, p. 300-305. (2015).
4. van der Weiden, A., Prikken, M., van Haren, N.E.M.: Self-other integration and distinction in schizophrenia: A theoretical analysis and a review of the evidence. Neuroscience and Biobehavioral Reviews. 57, 220237 (2015).
5. Weiss, C., Herwig, A., Schütz-Bosbach, S.: The Self in Social Interactions: Sensory Attenuation of Auditory Action Effects Is Stronger in Interactions with Others. PLoS ONE. 6, e227233 (2011).
6. Colonius, H., Diederich, A.: Multisensory Interaction in Saccadic Reaction Time: A Time-Window-of-Integration Model. Journal of Cognitive Neuroscience. 16, 10001009 (2004).
7. Kilner, J.M., Paulignan, Y., Blakemore, S.J.: An Interference Effect of Observed Biological Movement on Action. Current Biology. 13, 522525 (2003).
8. Kilner, J.M., Friston, K.J., Frith, C.D.: Predictive coding: an account of the mirror neuron system. Cognitive Processing. 8, 159166 (2007).
9. Pezzulo, G., Donnarumma, F., Dindo, H.: Human sensorimotor communication: a theory of signaling in online social interactions. PLoS ONE. 8, e79876 (2013).

Motor expertise facilitates the cognitive evaluation of body postures: An ERP study

Dirk Koester[1,2], Thomas Schack[1,2,3] & Iris Güldenpenning[4]

[1]Neurocognition and Action Research Group - Biomechanics; [2]CITEC; [3]CoR-lab, Bielefeld University; [4]Sport psychology, University of Paderborn

Human complex actions such as gross, whole body movements in sports (e.g., fosbury flop in high jump) are organised according to a (biomechanically) functional structure. The whole movement can be subdivided into several functionally bound sub movements. Memory representations for such movements have been shown to reflect these sub division (so-called basic action concepts; BACs) and have been related to a cognitive ("mental") level of representation (Schack, 2004; Schack et al., 2016); see Table 1.

Table 1: *The levels of the cognitive action architecture approach (Schack et al., 2016).*

Code	*Level*	*Main function*	*Content*
IV	Mental control	Regulation	Symbols, Strategies
III	Mental representation	Representation	Basic action concepts
II	Sensorimotor representation	Representation	Perceptual representations
I	Sensorimotor control	Regulation	Motor primitives, Basic reflexes

Obviously, temporal information is critical for successful and precise movement execution and should, thus, be contained within movement memory representation as it has been shown specifically for high jump movements (Güldenpenning et al., 2013). Also, the activation of temporal order information has been suggested to be an automatic process in athletes in an evaluation task regarding body postures of high jump movements (Güldenpenning et al., 2011). As movement representations (BACs) are stored at the level of mental representations, temporal order information should affect the cognitive evaluation of body postures.

We recorded the electroencephalogram (EEG) hypothesising that temporal order information affects the level of mental representation. The P300 component of the event-related potentials (ERPs) reflect cognitive evaluations of stimuli (context updating; Donchin and Coles, 1988; Polich, 2007). Thus, we expected a P300 modulation in athletes but not in novices if participants evaluate body postures of the fosbury flop. A subliminal priming paradigm was used in which various body postures of the approach or the flight phases were shown as prime and target reflecting the natural (prospective) or the reversed (retrospective) order of the movement phases. Both groups are hypothesised to differ qualitatively in processing because only athletes have an according motor programme at their disposal.

Method

A total of 33 right-handed sport students participated voluntarily; 17 novices (25,1 yrs.; 11 female) who had no practical experience with the high jump movement (or at most minimal experiences from school lessons) and 16 athletes with a focus on tack-and-field and specific experience in high jump (22,3 yrs.; 8 female). The athletes had acquired a sufficient motor representation for the high-jump movement as evidenced by practical performance. Participants gave written informed consent, and the study adhered to the ethical standards of the latest revision of the Declaration of Helsinki (Fortaleza; WMA, 2013).

Eight photographs from a high-jump movement recording (television broadcasting from the final contest of the Olympic Games 2008 in Peking) were used as stimuli (four approach and four flight phase images). Pre- and post-masks consisted of 25 × 25 randomised cut-outs (10 × 10 pixels) of the stimulus set, generated automatically by visually scrambling versions of the stimuli. Also, distraction from the irrelevant background was reduced by blurring. Stimuli were presented centrally, subtended a visual angle of 6.5° with a size of 9.0 × 9.0 cm (250 × 250 pixels). Stimulus presentation was controlled by Presentation® (version 14.1; http://www.neurobs.com). For the full stimulus set see Güldenpenning et al. (2011).

A 2 × 2 factorial design with the factors *congruency of movement phases* (same phase vs. different phases) and *temporal order* (prospective vs. retrospective) was employed in a subliminal priming paradigm. Participants were instructed to classify the target pictures as approach or flight image as fast and as accurately as possible via external push button-responses. The response button assignment was counterbalanced across subjects.

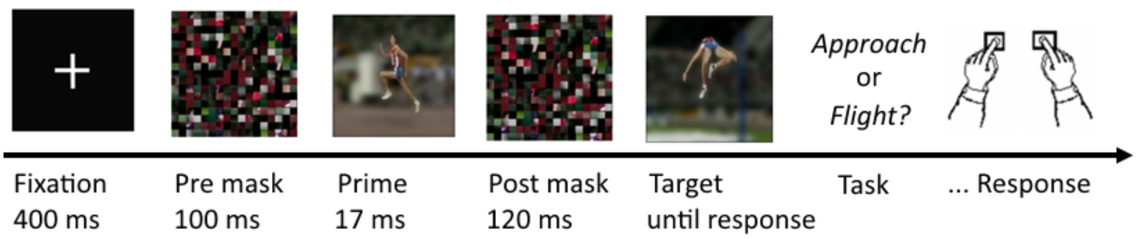


Figure 1: The trial procedure with prime and target examples. The inter trial interval was 1500 ms.

After twelve practice trials, participants performed 240 trials (24 prime target-pairs repeated ten times randomly) with a short break. For the trial timing see Figure 1. (Identical prime target pictures were excluded to avoid repetition priming.) Participants then performed a detection task on the primes (24 trials) to check the subliminal prime presentation.

The EEG was recorded (64 Ag/AgCl electrodes) based on the 10-10 system, low-pass filtered (DC-138 Hz) and sampled with 512 Hz. Eye movements were controlled for by recording the electrooculogram; impedances were kept below 5 kΩ. The EEG was band-pass filtered (0.1 - 30 Hz) and re-referenced to linked mastoids. Ocular artefacts were corrected using the Gratton algorithm. Automatic rejection was done using a moving window approach (200 ms extension; threshold ± 50 µV) and visually double-checked. The ERPs were time-locked to target onset with a 200 ms baseline before target onset. Average ERP amplitudes were calculated for each region-of-interest (ROI) and condition for the P300 time window (300-600 ms). The midline ROIs were frontal: FZ, FCZ; central: CZ, CPZ; posterior: PZ, POZ. Geenhouse-Geisser correction (ε) was applied were appropriate; corrected *p*-values, original *df* and effect size (Ω^2) are reported.

Results

Representative ERPs for both groups are shown in Figure 2. The ANOVA with the factors *temporal order, movement phase* and ROI (anterior vs. posterior, *AP*) for novices yielded no significant amplitude differences (all *F*s < 2.27; ns). The same ANOVA for the athletes yielded a main effect of temporal order ($F_{1,15}$=7.38; p<.05; Ω^2= 0.0383) which was qualified by an interaction of *temporal order* and *AP* ($F_{2,30}$=7.77; p<.01; ε=0.689 ; Ω^2= 0.0428). Separate *t*-tests for *temporal order* for frontal and central electrodes yielded no significant differences. Temporal order led to an increased P300 amplitude for prospective prime target pairs relative to retrospective picture pairs at posterior midline electrodes (t_{15}= 3.75; p<.01; Ω^2=0.1469). (Including *group* in the overall ANOVA as a between subject factor, led to a main effect of *group* ($F_{1,31}$=11.83; p<.01; Ω^2=0.0972) and an interaction of *group* and *AP* ($F_{1,62}$=4.35; p<.05; ε=0.830; Ω^2=0.00114) supporting qualitatively different processing.) Athletes' effect size for temporal order (d') differed from zero (0.353; posterior ROI; t_{15} = 3.75; p < .01; Ω^2=0.4011) but novices' did not (0.202; t_{16} < 1.42; ns). The peak latency did not differ among conditions (mean athletes: 442 ms; novices: 435 ms; all *F*s < 1; ns). The detection performance for primes was at chance level for novices (approach picture: 49.2 % correct; flight: 47.4 %) and athletes (approach: 48.4 %; flight 53.3 %).

Discussion

Supporting our hypothesis, *temporal order* elicited an increased P300 only in the group of athletes. The P300 amplitude was increased for prospective prime target-pairs at posterior electrodes in line with the natural order of the high jump movement (with a bilateral scalp distribution; not shown). The P300 effect suggests that the cognitive evaluation (classification) process was easier for prospective than for retrospective prime target picture pairings, possibly indicating higher response uncertainty for retrospective pairs (Horst et al., 1980). Thus, the P300 effect supports the theoretical ascription of movement representations to the level of mental representation (Schack 2004) in terms of neurophysiological processing.

Furthermore, the P300 effect was obtained with subliminal prime presentation. Hence, it is suggested that temporal order information has been processed automatically in athletes and may, therefore, be an integral part of the whole movement representation. Further research is needed to fully understand the (neurophysiological) functional relation between the mental and the action domains.

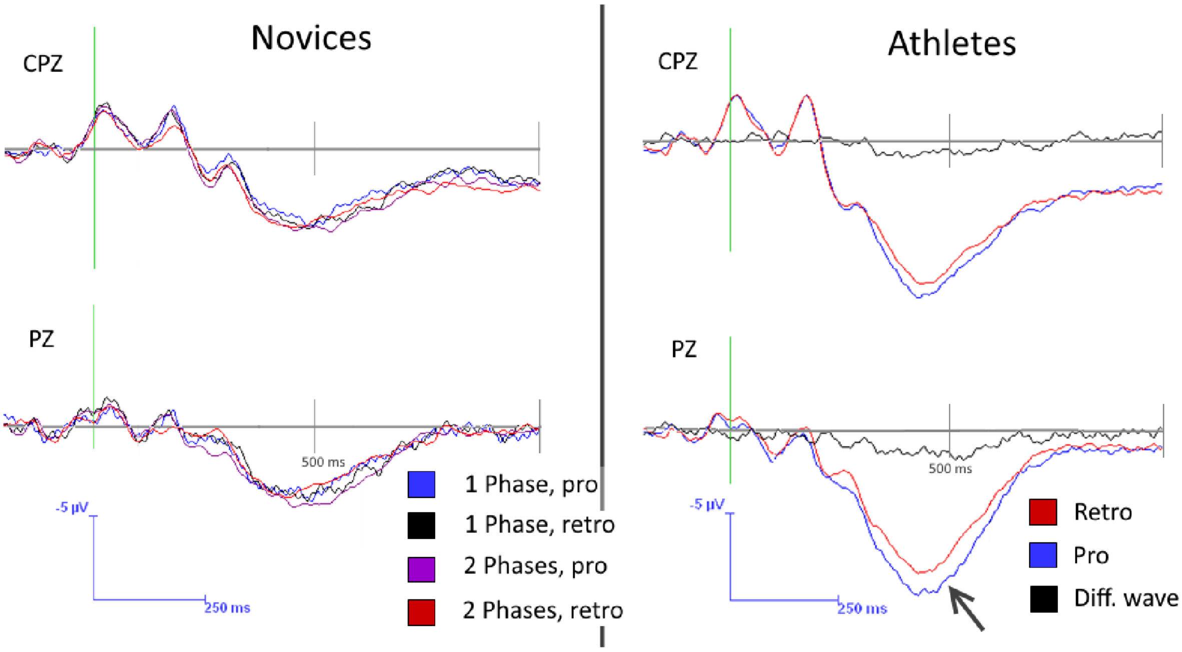


Figure 2: Grand average ERPs for novices (left panel) and athletes (right panel). Target onset at zero ms (vertical, green line; negativity up). Only athletes show an increased P300 amplitude for prospective (*Pro*) vs. retrospective conditions (*Retro*; 300-600 ms). Note, *1 Phase:* same-, *2 Phases:* different-movement phases.

References

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating. *Behavioral and Brain Sciences, 11*, 357-427.

Güldenpenning, I., Koester, D., Kunde, W., Weigelt, M., & Schack, T. (2011). Motor expertise modulates the unconscious processing of human body postures. *Experimental Brain Research, 213*, 383-391.

Güldenpenning, I., Kunde, W., Weigelt, M., & Schack, T. (2012). Priming of future states in complex motor skills. *Experimental psychology, 59*, 286-294.

Horst, R., Johnson, R., & Donchin, E. (1980). Event-related brain potentials and subjective probability in a learning task. *Memory & Cogntion, 8*, 476-488.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*, 2128-2148.

Schack, T. (2004). The cognitive architecture of complex movement. *International Journal of Sport and Exercise Psychology; Special Issue Part II: The Construction of Action - New Perspectives in Movement Science, 2*, 403-438.

Schack, T., Schütz, C., Krause, A. F., & Seegelke, C. (2016). Representation and anticipation in motor action. In M. Nadin (ed.), *Anticipation Across Disciplines*, Vol. 29 of Cognitive Systems Monographs (pp. 203-215). Springer.

Optimum statistical representation obtained from an intermediate feature level of the visual hierarchy.

Noshaba Cheema[1,2], Lex Fridman[3], Ruth Rosenholtz[3], and Christoph Zetzsche[1⋆]

[1] Cognitive Neuroinformatics, University of Bremen, Germany
[2] University of Applied Sciences Bremen, Germany
[3] CSAIL, Dept. of Brain and Cognitive Sciences, MIT, USA

Abstract. Representations obtained from the statistical pooling of features gain increasing popularity. The common assumption is that low-level features are best suited for such a statistical pooling. Here we investigate which level of a visual feature hierarchy can actually produce the optimal statistical representation. We make use of the award-winning VGG 19 deep network which showed human-like performance in recent visual recognition benchmarks. We demonstrate that the optimum statistical representation is not obtained with the early-level features, but with those of intermediate complexity. This could provide a new perspective for models of human vision, and could be of general relevance for statistical pooling approaches in computer vision and image processing.

1 Introduction

Representations that are based on a statistical pooling of features are of relevance in variety of contexts, ranging from bag-of-words models in natural language processing to texton approaches in computer vision [2, 5]. In human vision, the neural computation of *summary statistics* [1] (e.g. mean, variance, and cross-correlations of spatial filter responses) seems crucial for determining perception in the largest part of the field of view, i.e. for those 98% of total area being represented by peripheral vision and not by the high-performance central fovea [1, 3]. A common assumption about statistical pooling is that the features to be pooled should be relatively simple and low-level. In vision and image processing, for example, local wavelet-like features with different orientations and sizes are commonly utilised. However, convincing as this assumption may appear on a first look, it is actually far from clear why low-level features should be best suited for a statistical pooling. Rather, multi-level feature hierarchies, proven useful in various contexts, should be explicitly considered here. The visual system, for example, consists of a hierarchy of multiple subsequent processing stages in which the nature of the visual features becomes systematically more abstract, invariant, and general. Which level in such a hierarchy is actually best suited for a statistical pooling remains to be determined

⋆ corresponding author: zetzsche@informatik.uni-bremen.de

2 Methods

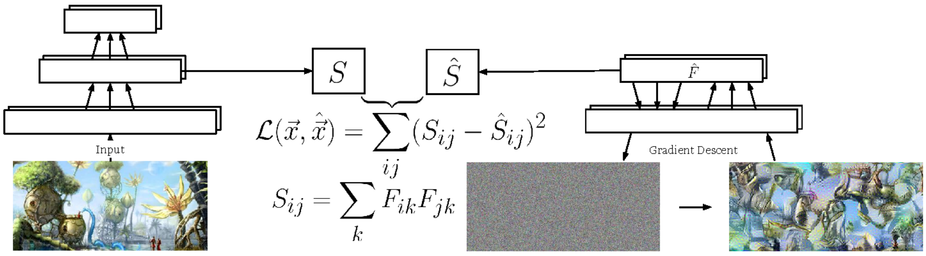


Fig. 1: Schematic architecture of the deep network and reconstruction algorithm. (F_{ik} : feature i at position k; S_{ij} : covariance of features F_i, F_j). Gradient descent is used with respect to the input image $\hat{\boldsymbol{x}}$ to minimize the loss $\mathcal{L}$ and to find a new image from white noise, which has the same statistics as the original.

Here we investigate this question by making use of an award-winning deep architecture, the 19 layer *VGG 19* net [6]. The impressive success of recent state-of-the-art *deep networks*, challenging for the first time human performance in visual recognition benchmarks [6], also indicates that such hierarchical architectures are crucial for obtaining efficient representations.The architecture of our model is shown in Fig. 1. The feature hierarchy is schematically illustrated on the left side. For the statistical pooling part (box S) we use a recent extension of this architecture for texture synthesis [4]. In texture synthesis, the goal is to generate a texture which has the same statistical properties as a given example texture. Here we make use of this process with a somewhat different conceptual attitude suggested by the Rosenholtz lab [1] by *reconstructing an individual image* from its statistical representation [1, 3]. In the original texture synthesis model, the statistical information is accumulated across layers, starting from the first layer and comprising all subsequent layers up to some specified maximum height in the hierarchy. We took a different approach here, and selected out only one single layer of the VGG-19 net in each test, since we wanted to know which layer of the hierarchy can provide the most valuable information for a statistical representation. This is schematically illustrated for the second layer in Fig. 1. We reconstruct the image from solely the statistical information S, which is provided by the covariance/spatial pooling of all respective features F_{ik}, F_{jk} of a layer l (i.e, we reconstruct from the *statistical representation* provided by the box S in Fig. 1). F is a feature map matrix, where k is the position of a feature and i the i^{th} filter. For comparison, we also applied the reconstruction procedure to the single "raw" features F of this layer (*feature representation*, leftmost side of Fig. 1). In other words, in the case of the "raw" *feature representation* we reconstruct directly from the features F_{ik} in a single layer of the hierarchy, whereas in case of the *statistical representation* we reconstruct from the covariance/spatial pooling of those features. We used 7 different test images.

3 Results

As expected, the best reconstruction from the "raw" feature representation is obtained using the initial, lowest level `conv1_1` of the network (left image, upper and middle row of Fig. 2). The reconstructions then become gradually more distorted if we proceed towards higher layers (from left to right in the upper two rows). This has to be expected, since the amount of information is systematically decreased by the increasing abstraction towards higher stages of the hierarchy. Also, as expected, the quality of reconstructions from the statistical representation is generally lower, due to the statistical pooling effects (lower row in Fig. 2). Suprisingly, however, we here observe no longer a monotonic decrease of reconstruction quality. Rather, the reconstruction quality is low at the initial stage `conv1_1`, then gradually increases up to the intermediate layers `pool2` and `pool3`, and finally deteriorates again for the higher layers in the hierarchy.

Fig. 2: Reconstructions from the "raw" feature representation (*upper row*, red framed square zoomed in *middle row*) and from the statistical representation (*lower row*). Hierarchical level increases from left to right. Reconstructions are obtained by using only the respective single layer of the hierarchy.

4 Discussion

For the statistical representation, the optimum reconstruction thus is not obtained at the initial layer (which provides the maximum amount of information) but at an intermediate level, where already a certain information loss has taken place. Nevertheless, the features of intermediate complexity seem to be better

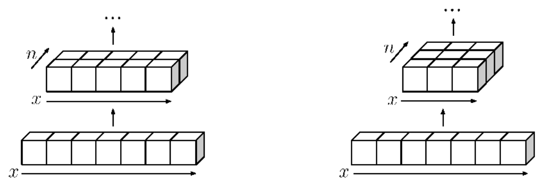


Fig. 3: Dimensionality reduction in different hierarchical architectures. Spatial resolution is decreased and feature number (and thereby the size of the statistical representation) is increased in both variants, but the tradeoff is different.

suited for a statistical pooling. This effect may be due to two counteracting factors. On the one hand, the total number of coefficients in the layers of the network is decreasing with increasing hierarchical level (a growing information loss) On the other hand the number of *different* features, and thereby the size of the statistical representation, and its associated differentation capability, is increasing. (Two example networks showing this effect are schematized in Fig. 3.) This trade-off may lead to an optimum at an intermediate stage. The detailed nature of the trade-off, however, depends on the specific architecture (cf. Fig. 3). For example, in preliminary tests with a network with large convolution kernels and a higher number of different features in the early layers, we observed a shift of the optimum towards the early stages. The true optimum for a statistical representation across *all* possible architectures thus remains to be determined. Furthermore, we have to extend the qualitative observations presented here by quantitative methods, e.g. by systematic tests of perceived similarity between original and reconstructed images, or by comparing clasification performance for reconstructed images from large test image databases [6]. In general, exploitation of the mutual dependence between the abstraction effect in feature hierarchies on the one hand, and the peculiarities of statistical measures and pooling effects on the other hand, can manage the information bottleneck in vision by providing qualitative information about the complete field of view.

References

1. Balas, B., Nakano, L., Rosenholtz, R.: A summary-statistic representation in peripheral vision explains visual crowding. Journal of vision 9(12), 13–13 (2009)
2. Cambria, E., White, B.: Jumping nlp curves: a review of natural language processing research. Computational Intelligence Magazine, IEEE 9(2), 48–57 (2014)
3. Freeman, J., Simoncelli, E.P.: Metamers of the ventral stream. Nature neuroscience 14(9), 1195–1201 (2011)
4. Gatys, L., Ecker, A.S., Bethge, M.: Texture synthesis using convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 262–270 (2015)
5. Renninger, L.W., Malik, J.: When is scene identification just texture recognition? Vision research 44(19), 2301–2311 (2004)
6. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

Auditory stimulus detection partially depends on visuospatial attentional resources

Basil Wahn[1], Supriya Murali[1], and Peter König[1,2]

[1]Neurobiopsychology, Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany
[2]Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
{bwahn,smurali,pkoenig}@uos.de

Keywords: Multisensory processing, Visuospatial attention, Multiple object tracking, Multisensory integration, Detection task

Abstract. Humans constantly receive and integrate information from different sensory modalities. However, the available attentional resources to process this information are limited. A matter of ongoing discussion in multisensory research is how attentional resources and multisensory processing are interrelated [1–3, 5–8, 11–14]. Recently, researchers suggested that the allocation of attentional resources in multisensory processing is task-dependent [11–14]. With regard to this task-dependency, it has not been investigated to what extent auditory detection task performance relies on visuospatial attentional resources. Moreover, it has not been investigated whether audiovisual integration in a detection task is affected when visuospatial attentional resources are diverted to a secondary task. Here, we addressed these two questions in a dual task paradigm ($N = 20$, $M = 22.6$ years, $SD = 3.09$ years, 14 female). In particular, participants performed a multiple object tracking ("MOT" [9]) task and a detection task either separately or at the same time. In the detection task (Figure 1A), participants either had to detect a white flash that always occurred within a black circle in the centre of the screen ("VI") or a "click" sound ("AU"). In a third detection task condition, both, the flash and click were presented simultaneously ("VIAU") and had to be detected. While participants performed the detection task, stimulus contrast (VI), loudness (AU), or both (VIAU) were adjusted depending on whether participants detected the stimulus or not using a QUEST staircase procedure [10]. In particular, as a dependent measure, we estimated for each condition the 75% detection task threshold separately. In the MOT task (Figure 1B), participants tracked a subset of several randomly moving objects. Here, performance was measured as the fraction of correctly selected targets. The experiment took about 2h. As the assumption of normality was frequently violated, we used non-parametric tests for the analysis.
For analyzing the performance with regard to the question whether attentional resources are shared or distinct, we calculated an overall score of task interference. For this score, we first calculated performance ratios for the MOT performance (see Figure 2A) by dividing the single task

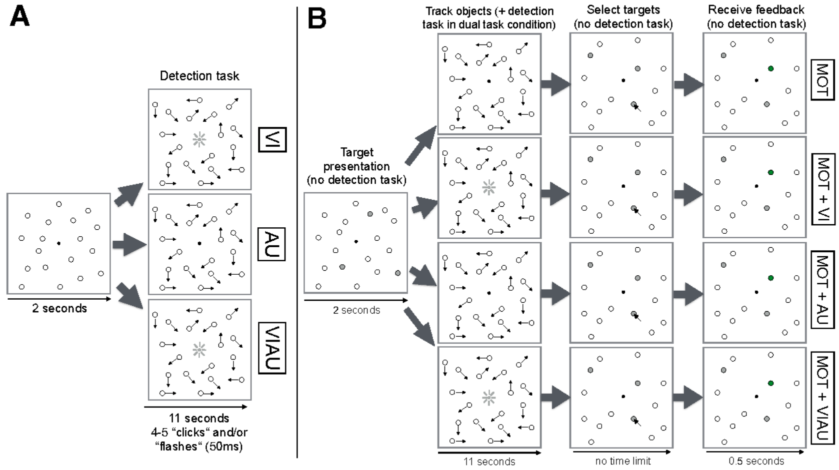


Fig. 1. A) Detection task overview. B) Multiple object tracking task overview. The trial logic is shown for performing the MOT task alone (top row), and in combination with either the visual (MOT+VI, second row), the auditory (MOT+AU, third row) or the audiovisual (MOT+VIAU, fourth row) detection task.

condition performance (MOT) separately by each of the dual task conditions (MOT+VI and MOT+AU). A value of above 1 indicates that the detection task interfered with performance in the MOT. With regard to the detection task performance (see Figure 2B), we divided the detection threshold obtained from the single task condition by the threshold in the dual task condition. To have a measure of the overall interference between tasks, we calculated the deviations from 1 for the detection task and MOT task ratios and added these deviations, separately for each participant and condition (see Figure 2C). We found that the interference was significantly above zero for VI and AU (VI: $W = 204$, $p < .001$; AU: $W = 174$, $p = .008$), suggesting that tasks interfered regardless of the sensory modalities in which they were performed. However, the overall interference between tasks was significantly smaller when tasks were performed in separate sensory modalities (i.e., vision and audition) than in the same sensory modality (i.e., vision) ($W = 161$, $p = .036$), suggesting that attentional resources are partly distinct.

For analyzing the performance with regard to the question whether audiovisual integration in a detection task is affected by diverting visuospatial resources, we divided the thresholds in the audiovisual detection task (VIAU) by the unisensory detection task thresholds (VI and AU). A value below 1 indicates that participants' perceptual sensitivities to detect stimuli benefited from receiving stimuli from two sensory modalities (i.e., vision and audition) compared to one sensory modality (either vision or audition). For the single as well as dual task condition, we found that participants' perceptual sensitivities were significantly below 1 (single: $W = 12$, $p < .001$, dual: $W = 19$, $p < .001$), suggesting

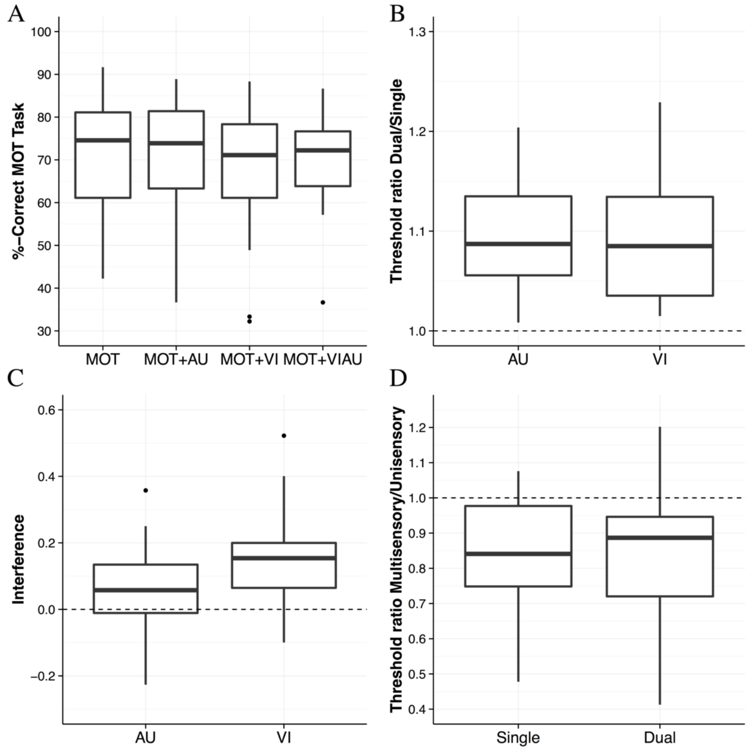


Fig. 2. Results overview. A) MOT performance (i.e., fraction correct of target selections) as a function of single task (MOT) and dual task conditions (MOT+AU and MOT+VI). B) Detection task ratios. C) Interference between the MOT and detection task. D) Multisensory detection task ratio. Box plots are shown in all panels (1.5 of the interquartile range is used for the whiskers).

that participants integrated the auditory and visual stimuli regardless of whether visuospatial attentional resources were diverted to the MOT task or not.

As a point of note, one might argue that the improved detection task ability when receiving audiovisual stimuli is not due to the process of multisensory integration but could be due to an alternation strategy. Specifically, participants could choose to only respond to the stimulus in the sensory modality that they can detect more easily in a given trial. However, this account seems unlikely given that detection thresholds were matched to equal sensitivity for each sensory modality prior to the experiment.

Given earlier studies [5, 11–14], the present study further supports the view that attentional resource allocation in multisensory processing is task-dependent and that multisensory integration for low-level stimuli is not dependent on attentional resources. Future studies could investigate whether this task-dependency generalizes also to other task combinations that have not been investigated yet [11–14]. Furthermore, findings are applicable to circumstances in which the visual sensory modality is already taxed with a demanding visuospatial task. In such circumstances, limitations in visuospatial attentional resources can be effectively circumvented by distributing information processing across several sensory modalities.

References

1. Alais, D., Morrone, C., Burr, D.: Separate attentional resources for vision and audition. Proceedings of the Royal Society of London B: Biological Sciences, 273(1592), 1339-1345 (2006)
2. Alsius, A., Navarra, J., Campbell, R., Soto-Faraco, S.: Audiovisual integration of speech falters under high attention demands. Current Biology, 15(9), 829-843 (2005)
3. Arrighi, R., Lunardi, R., Burr, D.: Vision and audition do not share attentional resources in sustained tasks. Frontiers in Psychology, 2:56, (2011)
4. Bertelson, P., Vroomen, J., De Gelder, B., Driver, J.: The ventriloquist effect does not depend on the direction of deliberate visual attention. Perception & Psychophysics, 62(2), 321-332 (2000)
5. Chan, J. S., Newell, F. N.: Behavioral evidence for task-dependent "what" versus "where" processing within and across modalities. Perception & Psychophysics, 70(1), 36-49 (2008)
6. Duncan, J., Martens, S., Ward, R.: Restricted attentional capacity within but not between sensory modalities. Nature, 387(19), 808-810 (1997)
7. Helbig, H. B., Ernst, M. O.: Visual-haptic cue weighting is independent of modality-specific attention. Journal of Vision, 8(1):21, 1-16 (2008)
8. Navarra, J., Alsius, A., Soto-Faraco, S., Spence, C.: Assessing the role of attention in the audiovisual integration of speech. Information Fusion, 11(1), 4-11 (2010)
9. Pylyshyn, Z. W., Storm, R. W.: Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision, 3(3), 179-197 (1988)
10. Watson, A. B., Pelli, D. G.: A Bayesian adaptive psychometric method. Perception & Psychophysics, 33(2), 113-120 (1983)
11. Wahn, B., König, P.: Vision and haptics share spatial attentional resources and visuotactile integration is not affected by high attentional load. Multisensory Research, 28(3-4), 371-392 (2015)
12. Wahn, B., König, P.: Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. Frontiers in Psychology, 6 (2015)
13. Wahn, B., Schwand, J., Krüger, M., Crafa, D., Nunnendorf, V., König, P.: Multisensory teamwork: Using a tactile or an auditory display to exchange gaze information improves performance in joint visual search. Ergonomics, 1-15 (2015)
14. Wahn, B., König, P.: Attentional resource allocation in visuotactile processing depends on the task, but optimal visuotactile integration does not depend on attentional resources. Frontiers in Integrative Neuroscience, 10 (2016)

Acknowledgments. We gratefully acknowledge the support by H2020 – H2020-FETPROACT-2014 641321 – socSMCs (for BW) and ERC-2010-AdG #269716 – MULTISENSE (for PK).

Recognition of self-performed, but visually unfamiliar dance-like actions from point-light displays

Bettina E. Bläsing

Neurocognition and Action Research Group and Center of Excellence
Cognitive Interaction Technology (CITEC), Bielefeld University
bettina.blaesing@uni-bielefeld.de

Keywords. Movement recognition · self recognition · dance-like actions · motor learning · point-light display

Abstract. Previous research has shown that motor experience of an action can facilitate the visual recognition of that action, even in the absence of visual experience. We conducted a point-light experiment in which participants were presented dance-like actions that were unfamiliar, visually familiar, familiar from visual learning and practice, or self-created while blindfolded. Accuracy of assigning the actions was above chance and equally high for all categories, including the self-created movements that the participants had never watched before. Self-recognition was above chance and was modulated by action experience, supporting the idea of an influence of agency on self-recognition.

1 Introduction

Performing or creating movement with closed eyes is considered a meaningful practice in dance training as it deviates from the common visual approach to movement learning. Creating novel movement only based on proprioceptive and kinesthetic feedback is expected to result in a multimodal mental movement representation that can be accessed via observation, despite the lack of visual familiarity of the movement. Casile and Giese [1] showed that motor experience of an action can facilitate the visual recognition of that action, even in the absence of visual learning or experience. The authors concluded that non-visual motor learning (based on haptic and verbal feedback) directly influences visual action recognition. The aim of the present study was to extend these findings to non-cyclic full-body movements that were not learned through haptic guidance or verbal instruction, but self-created without visual feedback. We investigated to what extent participants were able to visually recognize point-light displays of movements they had created and performed while they were blindfolded. Visual recognition of these movements was compared to movements that had been learned via observation and practice, movements that were only visually familiar, and unfamiliar movements. We expected that the self-created movements could be recognized successfully from observation equally well as movements that had been learned and/or observed, based on the multimodal nature of the action representation built up during movement creation. Furthermore, based on previous findings [2], we expected that participants would be able to distinguish between themselves and others as performers of the action.

2 Method

Participants: 19 right handed sports students (age 23.4±1.64 years, 4 males).

Procedure: The procedure consisted oft two sessions. During the first session, movement phrases were recorded with groups of students to produce point-light displays as stimuli for the following experimental session. In the second session, 14-21 days after the first, the students participated individually in an action recognition task.

Movement recording session: For the first session, students were randomly assigned to groups of three (in one case four). Two groups entered the biomechanics lab together for a recording session of approximately one hour, resulting in three recording sessions for all 19 participants. After entering the lab, one group was defined as „observers" and instructed to sit on the side of the lab, watching the other group attentively and quietly. The other group was defined as „active", and each member of this group was equipped with 15 retro-reflective markers (one on each foot, knee, hip, hand, elbow, shoulder; two on the forehead, one on the sternum). One member of the active group was blindfolded and guided to the recording space (approx. 2x2m). The two other members were standing outside the recording space, with sufficient space to move freely. The blindfolded participant was then instructed to start moving and create a short movement phrase that s/he considered novel and unusual, and to keep repeating it. The two partners were instructed to watch the and to learn the movement by observation and practice. After the two "learners" had indicated that they felt confident performing the movement, each member of the group was recorded performing the movement individually using a Vicon 3D motion capture system with 12 infra-red cameras. After the recording of this particular movement, the blindfold was removed from the eyes of the "movement designer", and one of the learners was assigned the new movement designer. The procedure was repeated until each member of the active group had once taken the role oft he blindfolded movement designer. Subsequently, the active group and the group of observers swapped roles, the observers were seated on the side and the procedure was repeated with the new active group.

Movement recognition task: The 3D recordings were transformed into 2D video clips, all viewed from the same distance and perspective (designated front view). Short clips each containing one full performance of a movement were cut from the footage to produce the stimulus material for the movement recognition task. For stimulus presentation, Presentation® software (14.8) was used. During the experiment, 36 movements were shown once in randomized order. Each movement clip was preceded by a screen indicating the running number of the following clip (2000ms), a black screen (500ms) and a fixation cross (500ms), and followed by a black response screen during which the stimulus presentation was paused until the participant pressed the space bar.

Each participant performed the recognition task individually in a quiet laboratory where s/he was presented 36 point-light clips showing 12 different movements, each performed by three different persons (the movement designer and two learners). Six of the movements (18 clips) were familiar (recorded during the session the participant had taken part in) and six were unfamiliar (recorded during sessions with other groups). For half of the familiar movements (9 clips), the participant had been observer, for the other half, s/he had been "active" (i.e., s/he had performed these movements him/herself, two sighted as learner and one blindfolded as movement designer). After each clip the stimulus presentation was paused, and the participant had to answer two questions (see Table 1) by ticking boxes in a paper questionnaire, and subsequently to press the space bar to activate the next stimulus presentation.

Table 1. Questions that had to be answered for each point-light clip during the experiment

Question 1: I have...	Question 2: The person in the video clip...
• neither watched nor performed this movement • only watched this movement, as observer • watched, learnt and performed this movement • created and performed, but not watched this movement	• was me • was not me

3 Results

Movement recognition: Data were categorized into correct answers (Q1: clips correctly assigned to movement categories; Q2: clips correctly identified as showing oneself or not oneself) and incorrect answers. For Q1, accuracy was calculated as proportion of correct responses from the total number of items of the category. Differences in accuracy were calculated using Wilcoxon signed-rank tests. Accuracy of recognizing the movement as unfamiliar, observed, learned or self-created (Q1) was higher than 0.8 for all categories, including the combined categories familiar (observed, learned, or self-created), visually experienced (observed or learned) and executed (learned or self-created) (see Table 2). Accuracy for all categories was above chance (0.25); no differences between any of the categories were found.

Table 2. Accuracy of assigning the presented point-light items (Question 1)

Category	Mean	SD
All items (n=36)	0.88	0.08
Unfamiliar (n=18)	0.93	0.08
Familiar (observed, learned, or self-created; n=18)	0.90	0.11
Visually experienced (observed or learned; n=15)	0.89	0.14
Observed (n=9)	0.82	0.20
Performed (learned or self-created; n=9)	0.87	0.15
Learned (n=6)	0.83	0.22
Self-created (n=3)	0.86	0.26

Actor recognition: Accuracy was calculated only for the performed (learned and self-created) movements, defined as combined accuracy of self- and non-self identification (proportion of correct responses from the total number of responses). Accuracy was 0.7±0.14 (self-created: 0.75±0.27, learned: 0.67±0.14). Participants identified relatively more actors as themselves (121%) and less as non-self (89%) when watching self-created movements, whereas they identified less actors as themselves (39%) and more as non-self (131%) when watching learned movements. A 2x2 repeated measures ANOVA was performed over the relative frequencies with factors ACTION (self-created, learned) and ACTOR (self, non-self), resulting in a main effect of ACTION [$F(0.896, 0.033)=27.377$; $p \leq .001$], a tendency for ACTOR [$F(1.521, 0.352)=4.318$; $p=0.052$] and an interaction [$F(6.810, 0.225)=30.272$; $p \leq .001$].

4 Discussion

Participants watched point-light displays showing dance-like actions of four categories: unfamiliar, observed (but not performed), learned (observed and performed) and blindfolded self-created (performed, but not observed). Based on previous studies [1], we expected that participants would be able to assign the presented actions correctly to the categories, independent of the modality of their previous action experience (visual, motor, both, or none). Accuracy of movement recognition was above chance and equally high for all categories. Crucially, this was also true for the self-created movements that the participants had never watched, corroborating that motor actions that one has performed but never visually observed can be recognized from visual observation. According to recent approaches, motor learning includes the integration of visual, verbal, proprioceptive and kinaesthetic information into a holistic multimodal mental action representation that underlies physical action execution as well as mental imagery, and its internal structure is reflected by the quality of action execution [3]. Results of the present study suggest that the mental representations of the learned and the self-created movements did not differ in quality, as both could be accessed via visual observation of the point-light displays.

It has been shown that people can easily distinguish between themselves and others when watching point-light displays performing motor actions [2], in particular dancing [4]. The observation that action experience (learned or self-created) obviously had an influence on actor identification extends these results and points towards an important role of agency for self-recognition [5].

The results of the presented study support a direct influence of motor experience on visual action perception and recognition for actions that have been learned without visual feedback. They extend previous results [1] to dance-like actions that have been acquired exclusively through movement exploration and practice, in the absence of vision and without haptic or verbal feedback. Furthermore, the results corroborate the position that action representation plays an important role for self-identification.

References

1. Casile, A., Giese, M.A.: Nonvisual motor training influences biological motion perception. Curr. Biol. 16, 69-74 (2006)
2. Loula, F., Prasad, S., Harber, K., Shiffrar, M.: Recognizing people from their movement. J. Exp. Psychol. Human 31, 210 (2005)
3. Land, W.M., Volchenkov, D., Bläsing, B., Schack, T.: From action representation to action execution: exploring the links between cognitive and biomechanical levels of motor control. Front. Comp. Neurosci. 7 (2013)
4. Sevdalis, V., Keller, P.E.: Self-recognition in the Perception of Actions Performed in Synchrony with Music. Ann. NY Acad. Sci. 1169, 499-502 (2009)
5. Jeannerod, M., Pacherie, E.: Agency, simulation and self-identification. Mind Lang. *19*, 113-146 (2004)

This is in memoriam of

Bruce Bridgemann

who was killed in an accident on 07 July 2016 in Taipei, Taiwan. The revised version of his contribution to KogWis 2016 included here was submitted by the author in mid June. We would like to recall the memory of Bruce Bridgeman and his contributions to cognitive science, deeply regretting his loss.

Space Constancy across Large Saccades

Bruce Bridgeman

Department of Psychology, University of California, Santa Cruz
Social Sciences 2 UCSC, Santa Cruz, Ca. 95064 ,USA
http://people.ucsc.edu/~bruceb/
bruceb@ucsc.edu

Abstract. Space constancy across large saccades is poorly handled by current theories. Observers made 36.5° saccades across a natural scene that never moved during data collection. Six cued control saccades were followed by asking whether the image had jumped during any of the saccades. If the answer was yes, observers were asked the jump's direction. Two of 28 observers reported jumps. After seeing the scene displace left and right, they were asked the same questions after each of ten more cued saccades. 68% of observers reported some apparent jumps (failures of constancy). Attention to possible jumps broke space constancy, implying that inattention mediates constancy for large saccades. The rules for achieving space constancy are qualitatively different for large saccades than for small ones.

Keywords: space constancy, saccades, reference object theory, efference copy

1 Introduction

Space constancy, the perception that the visual world remains stable despite retinal image displacements with each saccadic eye movement, is the platform on which all subsequent spatial processing relies, along with vestibular and other sensory signals.

For more than a century the dominant theory for space constancy invoked an efference copy or corollary discharge [1, 2], a motor signal to the extraocular muscles accompanied by a copy that informs the perceptual system of the anticipated retinal shift. The efference copy is subtracted from the retinal shift, resulting in constancy.

Empirical measures of the efference copy, however, found that its gain is too low to support constancy [3, 4] and its dynamics are too slow to keep up with saccadic eye movements [5, 6]. Efference copy theory also predicts that displacing the entire visual world during a saccade should be easily detected; the copy and the corresponding retinal shift would no longer match. This prediction too was contradicted [7]. The threshold displacement (50% of displacements detected) is nearly 1/3 as large as the saccade [3]. These results have led to theories [8] that reject compensation, instead holding that spatial values are calculated anew after of each saccade. First the visual system chooses a target object. If that object lands in or near the foveal region at restoration of fixation, previous spatial values are discarded [9] and the world is perceived as stable, a 'reference object theory' [10].

Previous studies of space constancy have been limited to small saccades, usually 10deg or less. The following experiment tests the role of inattention in achieving space constancy across large saccadic eye movements.

2 An Experimental Test

2.1 Observers

There were 28 observers, 8 male, undergraduates at the University of California, Santa Cruz. They were naïve to the purposes of the experiment.

2.2 Apparatus

Observers sat before a tangent screen 1.62m wide x 2.33m high at a distance of 1m. On the screen an image 49° wide x 38.5° high was projected through a front-surface mirror, a photograph of a natural forest scene with a distant ridge flanked by two trees with trunks extending above the horizon. The trees, 36.5° apart, served as fixation and saccade targets. The projected image was generated by a computer with software that could cause the image to jump laterally by 34 min arc (1cm) with single-keystroke commands.

Observers wore an eye monitor mounted on a modified spectacle frame, with an eye camera and an infrared LED positioned on a stalk extending to the lower right of the right eye. Between the eyes was a scene camera and the associated electronics. An eye monitor controller console on a table beside the observer was not compatible with the camera-based sensors that the observer wore. The large saccades were verified by direct observation of the observers' eyes.

2.3 Procedure

During data collection the image never moved. First the observer made 6 practice eye movements alternating leftward and rightward. Then in a calibration phase observers made alternating eye movements from one fixation target to the other on the experimenter's cue, each followed by the experimenter manipulating switches and dials on the eye monitor controller. These adjustments did not affect the progress of the experiment. The experimenter then stated "Now we're going to measure the dynamics of your eye movements." Then 6 saccades were executed on the experimenter's cue, alternating left and right, with at least 1sec between saccades. This constituted a control condition, with no reference to space constancy or the possibility that the image might jump or displace during a saccade.

At this point a question was posed, 'Did the image ever seem to jump at the same time as your eye movements or not?' Both possibilities were mentioned to minimize demand characteristics. If the answer was yes, the observer was asked 'Was the displacement in the same direction as the eye movement or in the opposite direction?'

The question directed the observers' attention toward the possibility of displacement. Demand characteristics were minimized by giving two alternatives for each question and never mentioning the attentional hypothesis of the experiment.

After this phase the experimenter pressed keys on the computer keyboard to cause square-wave displacement of the image to the right and left, explained as 'zeroing the display'. Finally, 10 saccdes were executed, again each at the experimenter's cue, and the questions above were posed after each one. The image never moved during these saccades. Saccades were cued one at a time to prevent fatigue. The questions served both to assess perceived motion and to direct the observers' attention.

3 Breakdown of Space Constancy

In the initial control phase only 2 observers (7%) reported displacement. In the experimental condition space constancy failures measured as judgments that the target had moved during a saccade occurred at least once for 61% of observers. Thus attention to possible displacements, elicited by the space-constancy questions, interfered with space constancy with no other changes in conditions. Observers varied in the frequency of reporting displacement; 2 always saw displacement, while 8 never did; the remaining observers saw displacement on some trials. The observers who always saw displacement in the experimental trials did not report any displacement in the control trials, even though stimulus conditions were identical in the control and experimental conditions. Overall, displacement was perceived on 28% of the 230 experimental trials. In terms of signal detection theory all of these trials were false alarms.

There was no consistent pattern of the frequency of perceived displacements as trials progressed, either for displacements with or agains the direction of the corresponding saccade (Figure 1).

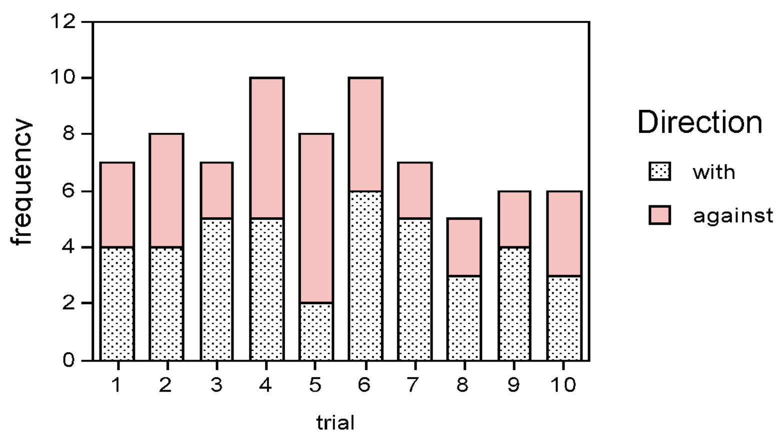


Fig. 1. Displacement judgements with and against saccade direction by trial.

In previous work with smaller saccades, breakdown in constancy normally occured with displacement detected in the direction opposite the saccade. Our observers, however, reported displacement in the same direction as the saccade in 55.8% of the trials in which displacement was seen, not significantly different from the opposite-direction trials (p=0.36, binomial sign test). The reference object theory [10, 11] can account for this difference. Since the theory requires foveal or near-foveal acquisition of the reference object to achieve space constancy following a saccade, undershoot should result in a failure of constancy. Because many large saccades undershoot, the post-saccade reference object will often appear peripherally in the saccade direction. Thus the visual world will appear to jump in the direction of the reference object.

Our result seems to conflict with earlier observations that real jumps of the visual world during small saccades are not detected. The conflict can be reconciled with a combination of range nonlinearity and attentional manipulation. For the large saccades used here, causing the observers to attend to possible target displacements caused some of the large saccade-associated retinal displacements to be interpreted in part as jumps of the image, a breakdown in space constancy from attention alone. Inattention thus aids in maintaining space constancy across large saccades.

4 References

1. Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43,* 482–489.
2. von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip: Wechselwirkungen zwischen Zerntalnervensystem und Peripherie. *Naturwissenschaften, 27,* 464–476.
3. Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research, 15,* 719–722.
4. Bridgeman, B. (2007). Efference copy and its limitations. *Computers in Biology and Medicine 37,* 924-929.
5. Grüsser, H. J., Krizic, A., & Weiss, L. R. (1987). Afterimage movement during saccades in the dark. *Vision Research, 27,* 215-226.
6. Ilg, U., Bridgeman, B., & Hoffman, K.-P. (1989). Influence of mechanical disturbance on oculomotor behavior. *Vision Research, 29,* 545-551.
7. Mack, A. (1970). An investigation of the relationship between eye and retinal image movement in the perception of movement, *Perception & Psychophysics, 8,* 291–298.
8. Bridgeman, B., van der Heijden, A. H. C., & Velichkovsky, B. (1994). Visual stability and saccadic eye movements. *Behavioral and Brain Sciences, 17,* 247–258.
9. McConkie, G., & Currie, C. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception & Performance, 22*, 563-581.
10. Deubel, H., Bridgeman, B., & Schneider, W. X. (2004). Different effects of eyelid blinks and target blanking on saccadic suppression of displacement. *Perception and Psychophysics, 66*, 772–778.
11. Deubel, H., Koch, C., & Bridgeman, B. (2010). Landmarks facilitate visual space constancy across saccades and during fixation. *Vision Research, 50,* 249-259.

Modeling Kitchen Knowledge with LTM^C

Marc Halbrügge[1] and Holger Schultheis[2]

[1] Quality & Usability Lab, Technische Universität Berlin
Ernst-Reuter-Platz 7, 10587 Berlin
marc.halbruegge@tu-berlin.de
[2] Cognitive Systems, Universität Bremen
schulth@informatik.uni-bremen.de

Abstract. The integration of knowledge about the world is an open problem of computational cognitive modeling. We are approaching this issue in the context of sequential control. Based on the Memory for Goals theory, we hypothesize that subgoals that are semantically related receive mutual priming. This hypothesis is tested using error data from a household scenario. We use LTM^C to represent facts about cooking recipes (e.g., "Tiramisu is a dessert") and to compute activation spreading between them. The resulting activation patterns are applied to action sequences that partially relate to some of these facts. A computer simulation of a corresponding cognitive model yields omission rates that fit well to empirical data.

Keywords: Spreading Activation; Long-Term Memory; Human-Computer Interaction; Memory for Goals

1 Introduction

Understanding human error is crucial for understanding how humans organize actions to attain their goals. A promising model of the sequential control of actions is the Memory for Goals (MFG) theory [1]. The central assumption of the MFG is that goals that correspond to individual actions are held in memory, thereby being subject to activation decay, priming and interference. Computational models based on the MFG have been shown to explain important facets of procedural performance and errors in laboratory and applied domains (e.g., [1, 6]). In this paper, we are exploring the advantages of combining semantic priming with the activation-based approach of the MFG theory.

1.1 Modeling Sequential Control and Error

Our cognitive error model extends on the MFG by highlighting the importance of external cues for sequential action [6]. The MFG has been originally validated using memory based planning tasks like the Tower of Hanoi [1]. Even when applied to problems in human-computer interaction, most existing MFG models keep all task knowledge in memory, comparable to top-down (vertical) processing in the model proposed by Cooper and Shallice [3]. As we have shown in [6], this purely memory based view can break down when it is applied to real-world applications, e.g., when some interactive elements of the user interface (UI) are optional, but others are obligatory. While our data (and daily life) show that obligatory tasks (e.g., logging into a computer system) are less prone to omissions than non-obligatory ones, the MFG lacks a straightforward explanation for this difference. We therefore proposed an additional process that takes over as soon as memory gets weak (comparable to horizontal triggers in the model by Cooper and Shallice [3]). This process scans the UI for interactive elements and uses a memory retrieval heuristic to check whether an element was part of the current action sequence. The addition of such a visual cue-seeking process does not only account for the difference between obligatory and optional tasks, it has also been shown to allow explaining intrusion errors that were beyond the reach of previous MFG models [6]. A recent eye-tracking study adds further evidence that world-based processing is an important aspect of sequential control [7].

Another important weakness of current models of sequential behavior is the neglection of pre-existing knowledge when new action sequences are formed and executed. In a previous analysis, we have linked error rates and task execution times in a tag-based search paradigm to the number of mentions of the respective search tag (e.g., "German") in a Wikipedia-based ontology [5]. While the results were promising, the approach lacked the possibility of associative priming *between* the subgoals in a longer task sequence. This kind of priming was beyond the technical capabilities of the cognitive architecture used there (ACT-R; [2]). For the current study, we therefore recreated the cognitive processes of the error model in a new system that uses an alternative model of human long-term memory, LTM^C, for knowledge representation.

1.2 LTM^C

LTM^C [8] represents knowledge as a network of nodes. Each node stands for an object or a relation. Each node comprises a unique identifier (e.g., the name of the represented entity) and an activation value, which determines the node's availability. Nodes are connected by links such that a link between two nodes indicates that the two entities represented by the nodes are associated to each other. A node's activation is determined from three main influences: base level activation, noise, and spreading activation. Of main importance for the work presented here is spreading activation, which distributes activation from the task context to the nodes in LTM^C via associative links: The more strongly a node is associated to the current task context, the higher its activation. LTM^C has been proven to provide an accurate account of human memory also when dealing which rich, real-world knowledge bases [8, 9].

2 Error Model with World Knowledge

We explored the implications of semantic priming on human error based on data that had been collected during an experiment with a kitchen assistance system (Figure 1, left). Twenty participants solved tasks like "Search for German main dishes and select Sauerbraten" or "create an ingredients list for three servings and check off salt and flour" (detailed materials and procedure in [6]). For the current analysis, we transformed

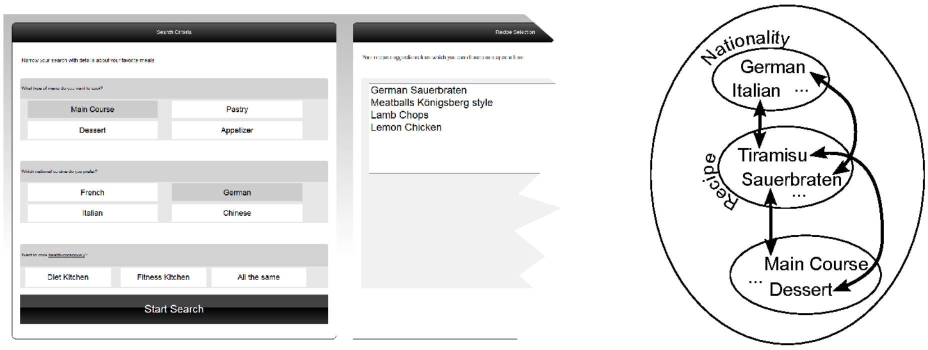


Fig. 1. English version of the kitchen assistant on the left; representation within LTM^C on the right. Circles and text denote objects within their categories, arrows denote relationships. Activation spreads along the arrows.

the semantic mappings between recipes and search attributes to an LTM^C object-relation-graph (Figure 1, right). An existing user interaction modeling system [4] was used to represent the action sequences of the experimental procedure as chains of subgoals (e.g., "German", "Sauerbraten"). Each subgoal was augmented with a numerical activation value. LTM^C was then used to modulate these activation values through activation spreading, i.e., in terms of semantic priming we treated the retrieval of the subgoal "Sauerbraten" similar to the retrieval of the generic "Sauerbraten" chunk from long-term memory. If the resulting activation was below a fixed threshold, the retrieval of the subgoal failed and the corresponding action was not carried out. The model's omission rates for different elements of the UI are given in Figure 2 (100 model runs); the goodness-of-fit is satisfactory with R^2=.73 and RMSE=.0073.

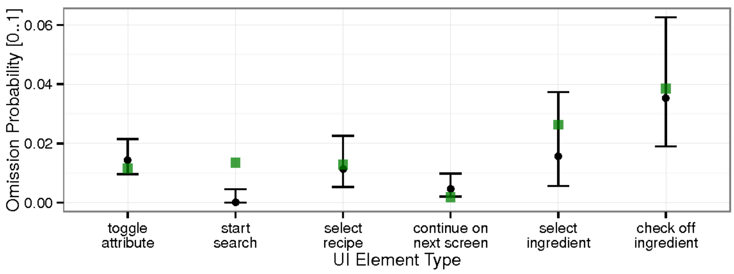


Fig. 2. Omission probabilities for different UI elements. ■ denote model predictions, error bars are 95 % confidence intervals using the Agresti-Coull method.

Especially noteworthy are the low model omission rates for 'toggle attribute' and 'select recipe'. Without semantic priming, these predictions would be on the level of selecting and checking off ingredients. The low omission rates for 'start search' and 'next screen' on the other hand are due to visual priming during cue-seeking.

3 Discussion

We have presented a model of sequential action in a household environment. By modulating the activation of individual subgoals based on semantic priming within LTM^C, we achieved a good fit to previously recorded data.

In principle, these averaged predictions could have been achieved using the Wikipedia-based approach presented in [5], but using LTM^C has several advantages. First, LTM^C's assumptions seem to better fit the mechanisms underlying human semantic priming, because we needed to completely override ACT-R's activation mechanism to achieve similar effects in [5]. Second, the priming in [5] was static, i.e., a subgoal 'German' received the same amount of additional priming regardless of the current task context. As the experimental procedure also contained tasks like "Switch from German to Italian and select Tiramisu", this approach led to the improbable result of 'German' being strongly primed during the 'Tiramisu' trial. This does not happen using the current solution (see arrows in Figure 1).

The validity of our approach is nevertheless limited. Because LTM^C is only containing knowledge extracted from the application (instead of an external ontology), the priming results sometimes do not match our participants' conceptions. An example is the 'Ratatouille' recipe which is flagged as 'main dish' in the kitchen assistant. Several participants strongly objected this view during our experiment.

Acknowledgments This work was partially supported by DFG grant MO 1038/18-1 ("Automatische Usability-Evaluierung modellbasierter Interaktionssysteme für Ambient Assisted Living").

References

1. Altmann, E.M., Trafton, J.G.: Memory for goals: An activation-based model. Cognitive science 26(1), 39–83 (2002)
2. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychological review 111(4), 1036–1060 (2004)
3. Cooper, R., Shallice, T.: Contention scheduling and the control of routine activities. Cognitive Neuropsychology 17(4), 297–338 (2000)
4. Engelbrecht, K.P., Kruppa, M., Möller, S., Quade, M.: MeMo workbench for semi-automated usability testing. In: INTERSPEECH. pp. 1662–1665 (2008)
5. Halbrügge, M., Quade, M., Engelbrecht, K.P.: How can cognitive modeling benefit from ontologies? Evidence from the HCI domain. In: Bieger, J., Goertzel, B., Potapov, A. (eds.) Proc. AGI'15, LNAI, vol. 9205, pp. 261–271. Springer, Berlin (2015)
6. Halbrügge, M., Quade, M., Engelbrecht, K.P.: A predictive model of human error based on user interface development models and a cognitive architecture. In: Taatgen, N.A., van Vugt, M.K., Borst, J.P., Mehlhorn, K. (eds.) Proc. ICCM'15. pp. 238–243. University of Groningen, Groningen, NL (2015)
7. Halbrügge, M., Quade, M., Engelbrecht, K.P.: Cognitive strategies in HCI and their implications on user error. In: Proc. CogSci'16 (2016), accepted
8. Schultheis, H., Barkowsky, T., Bertel, S.: LTM C – an improved long-term memory for cognitive architectures. In: Proc. ICCM'06. pp. 274–279 (2006)
9. Schultheis, H., Lile, S., Barkowsky, T.: Extending ACT-R's memory capabilities. In: Proc. EuroCogSci'07. pp. 758–763. Lawrence Erlbaum Associates (2007)

Model-driven Interaction Strategies of a Dialog System for Navigation and Information

Felix Putze, Elias Bordolo, and Tanja Schultz

University of Bremen, Cognitive Systems Lab,
Enrique-Schmidt-Straße 5, 28359 Bremen, Germany
{felix.putze,tanja.schultz}@uni-bremen.de
http://csl.uni-bremen.de

Abstract. In this work, we describe how a cognitive user model can be used in a spoken dialog system to determine the dialog strategy and the level of initiative.

Keywords: spoken dialog system; cognitive user model; memory model; initiative level

1 Introduction

The behavior of a spoken dialog system (e.g. a virtual call center agent or a voice-controlled in-car information system) is determined by its dialog strategy. The strategy can be understood as a mapping of the dialog state to an action of the dialog system, for example the generation of an utterance (the dialog state comprises information on the exchanged utterances, the task, the user, etc.). Thus, the dialog strategy determines the behavior of the system. One of the central aspects of system behavior is the level of initiative, i.e. the degree to which the system controls the dialog flow. Traditionally, researchers have discriminated between system initiative (SI) and user initiative (UI; often also mixed initiative). Earlier work based the choice for intiative level on user expertise and dialog success [2]. Newer research shows that the optimal level of initiative also depends on the user's workload level [6]. In this work, we propose to use a cognitive user model as a center piece of the dialog state which captures the user's state of memory and infers the user's interests. We employ this model to drive the dialog strategy and to provide a generalizable method to influence the initiative level of a system gradually. We analyze the proposed approach in a user study, in which we investigated two hypotheses regarding the use of model-based interaction strategies: 1) a model-driven interaction strategy can be used to select appropriate information at appropriate times, and 2) the parameters of the model-driven interaction can be configured to generate noticeably different levels of initiative. We investigate those hypotheses for a dialog system which is applied in a car-driving situation: The driver receives routing information as well as information on points-of-interest (POIs) along the route, transmitted via synthesized speech. The system can also request information on the user's interest or the user may give this information at their discretion. For modeling this application domain, we implemented an abstract representation of this scenario, in which simple graphical and textual cues are given to the user to describe the scene. As we are especially interested in investigating high-workload situations, the main driving task is complemented by a secondary task, following the n-back paradigm.

2 User Model & Interaction Strategy

The task of our cognitive user model is to estimate which pieces of information are most relevant to the user in the given context and to quantify the expected gain from additional information given to or requested from the user. This is challenging as the decision is influenced by the presence and timing of other relevant and irrelevant information. The estimate of the model will then be used by the system to determine whether to give (request) information from (by) the user, which information to give (request) and when to give (request) that information. By modifying the parameters of this mapping from user model to system utterances, we can influence the level of initiative of the system. The model-driven interaction strategy consists of three main components: First, an interest model which reflects the persistent aspects of relevance of information. Second, a memory model which reflects the volatile aspects of relevance of information. Third, a set of decision rules to derive system behavior from the two models.

The interest model captures the persistent aspects of relevance of information. We assume that every user has a fixed number of topics he or she is interested in. Topics are organized in an ontology to allow aggregation of interests (e.g. "museum" and "theater" are both child nodes of "culture"). Initially, the system is not aware of those interests. It collects information on those interests from questions it asks the user and from spontaneous user input to the system. The goal of the interest model is to integrate those individual pieces of information. For this purpose, it is implemented as a Bayesian Network representing the likelihood of each topic to be of interest for the user. The topology of the network represents the ontology of topics. Conditional probabilities in the modes are set to model a fuzzy-or between the child nodes. Received information on those interests is set as evidence in the corresponding nodes of the network and likelihood is inferred for other topics.

The memory model captures the volatile aspects of relevance of information, i.e. which information is available to the user. It reflects that recent information is readily available and decays over time, but is also influenced by the presence of other memory items competing for limited memory capacity. The memory model is based on the ACT-R theory [1] and its variants proposed in [5] and [3]. For each POI and each associated information slot (we refer to both as *concepts*), we define *activation* as the degree of availability of that concept. To calculate the activation of an item, we use the base-level activation mechanism (BLA) of ACT-R which reflects the impact of frequency and recency of stimulations. The decay process ensures that concepts which were stimulated in the past and which were not rehearsed will exhibit diminishing activation. BLA itself is not competitive between different concepts. Therefore, to enforce a soft memory capacity limit, we extended the model with an inhibition component: To calculate final activation, BLA of each concept is scaled with a sigmoidal threshold function of the relative BLA. The effect of this procedure is that concepts with high relative base-level activation remain mostly unchanged while the activation of other concepts is diminished. From final activation, we calculate the retrieval probability as suggested by the ACT-R theory.

The decision rules of the interaction strategy define how the model information is translated to system behavior. There are three main decision rules which control which utterance is generated: 1) Give route information to the user, 2) Give POI information to the user, and 3) Ask user about interests. Each decision rule checks a condition to determine whether the rule is triggered. If no decision rule is triggered, the system stays quiet until the next evaluation step.

For defining the trigger conditions of the three main decision rules, we need to introduce two concepts derived from memory model and interest model: information gain and uncertainty reduction. For a potential route or POI information (regarding concept c_i), we calculate the *information gain* from the memory model. The information gain measures the potential increase in predicted retrieval probability of c_i under the assumption that c_i is given to the user at the current time (weighted by the estimated interest in c_i). Thus, information gain depends on the current activation level of c_i, the estimated interest in c_i and the estimated duration until c_i must potentially be retrieved. For an interest question, we calculate the potential *uncertainty reduction* of the Bayesian network, measured by the decrease of entropy for all nodes. We calculate the uncertainty reduction in case of a "yes" and a "no" response separately and average the results, weighted according to the probability of each answer.

With those definitions in place, we can formulate the exact decision rules for POI information, route information and interest questions: For the first two types, we define thresholds t_{poi} and t_r. If the maximum information gain for presenting a POI (route) information is above t_{poi} (t_r), the corresponding utterance is generated. We proceed analogously for interest questions and the uncertainty reduction threshold t_u. By choosing different values for t_{poi}, t_r, and t_u, we can now define different initiative levels for the dialog strategy. When values are comparably low, the system will generate its utterances pro-actively in more situations, which corresponds to a system initiative behavior. When values are comparably high, the system follows a user initiative behavior. Beyond a dichotomized understanding of initiative, this approaches also allows a gradual transition between SI and UI.

3 Evaluation

To evaluate the model-based dialog strategy, we conducted a user study with 10 university students as participants. During the study, participants went through two different instances of the task, one time with SI and one time with UI behavior (the order of initiative strategies was randomly counterbalanced between participants). The two task instances used different content but were structurally identical (e.g. regarding timing and duration of road segments) for comparison. During each instance, multiple quizzes asked the user for specific POI or route information. Besides objective interaction parameters, we also handed out a questionnaire after each task instance.

First, we investigated the quality of the responses given by the participants during the quizzes. For the POI quizzes, participants achieved a correctness rate of 58.2% in SI and 57.7% in UI. For the route quizzes, participants achieved correctness rates of 78.8% and 71.2%, respectively. As participants always had to chose between three options, those results show that participants were able to answer quizzes significantly better than random guessing. Between initiative levels, there were differences in how the information was presented: In SI, POI information was given earlier (on average, POI information was given 62.7s before a quiz for SI, compared to 44.2s for UI) and more un-interesting information was given (on average, 3.0 uninteresting infos for SI between two quizzes, compared to 1.8 for UI). Furthermore, participants reported that they made mistakes during quizzes due to forgetting of information in SI, but due to omissions of information by the system. The questionnaire shows that for UI, user report to take the initiative, to wait less for system activity, to be less frustrated and to have a higher level of control over the information flow than for SI. The subjective assessment is also supported by the objective statistics on the distribution of speech acts: For the SI strategy, the system generates 49.3% of all utterances in a dialog; for the UI strategy, this decreases to 36.4%. The differences between SI and UI are in line with the observations from Putze and Schultz [4], showing that a system taking initiative leads to increased task performance but a subjective feeling of intrusiveness and loss of control. In summary, we see that a model-based interaction strategy is able to give appropriate information to the user and can be employed to generate different levels of initiative.

References

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychological Review 111(4), 1036–1060 (2004)
2. Litman, D.J., Kearns, M.S., Singh, S., Walker, M.A.: Automatic optimization of dialogue management. In: Proceedings of the 18th Conference on Computational Linguistics. pp. 502–508 (2000)
3. Pröpper, R., Putze, F., Schultz, T.: JAM: Java-based associative memory. In: Proceedings of the Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop. pp. 143–155 (2011)
4. Putze, F., Schultz, T.: Investigating intrusiveness of workload adaptation. In: Proceedings of International Conference on Multimodal Interfaces (2014)
5. Schultheis, H., Lile, S.: Extending ACT-r's memory capabilities. In: Proceedings of the Second European Cognitive Science Conference (2007)
6. Vetek, A., Lemmelä, S.: Could a dialog save your life?: analyzing the effects of speech interaction strategies while driving. In: Proceedings of the 13th international conference on multimodal interfaces (2011)

What are the contents of representations in predictive processing?

Wanja Wiese (wawiese@uni-mainz.de)

Johannes Gutenberg University
Department of Philosophy
55099 Mainz, Germany

We are currently witnessing a surge of research on predictive processing (PP), both in cognitive science and philosophy (Clark, 2016; Hohwy, 2013). The framework's kinship with the Bayesian brain hypothesis (Vilares & Kording, 2011) and the use of forward models (which it shares with optimal control approaches, see Pickering & Clark, 2014) provoke the question whether PP offers any unique explanations of cognitive phenomena. More specifically, we can ask if PP has a unique story to tell about the computational underpinnings of mental representations (cf. Wiese, 2016).

This meta-theoretical paper explores the contents of representations in PP (thereby extending results from Gładziejewski, 2016), by applying Frances Egan's distinction between mathematical and cognitive contents (Egan, 2014). The goal is to specify what relevance subpersonally ascribed, computational states (like Bayesian "beliefs") in PP have for personal-level descriptions (of full-blown beliefs and conscious states).

A first result is that all PP models are committed to the following core mathematical contents: estimates (first- order statistics), predictions, prediction errors, and precision estimates (second-order statistics). As such, they do not put any unique constraints on ascriptions of cognitive contents (e.g., contents of beliefs or conscious states). However, the hierarchical nature of processing in PP, in which the interplay between (first-order) estimates at different levels is mediated by differing precision estimates (Clark, 2013), provides specific constraints on explanations of, among others:

- attentional phenomena (Feldman & Friston, 2010),
- symptoms of mental disorders (Friston, Stephan, Montague, & Dolan, 2014),
- time perception (Arnal & Giraud, 2012; Hohwy, Paton, & Palmer, 2015)
- emotions (Seth, 2013), and
- motor control (Adams, Shipp, & Friston, 2013), to name just a few.

Additional unique constraints are provided by particular PP models, e.g. by models developed within the framework provided by Karl Friston's free energy principle Friston (2010). More specifically, in models using the Laplace approximation, important estimators are the sufficient statistics of the normal distribution (sample mean and covariance, Friston, Daunizeau, Kilner, & Kiebel, 2010, p. 234). To the extent that mean values constitute central aspects of computational models of cognitive phenomena, they put specific constraints on ascriptions of cognitive contents, since they strongly support the assumption that many perceptual contents are indeterminate (Madary, 2012), just like summary statistics (which are indeterminate with respect to individual samples). This idea can be applied to a variety of perceptual phenomena, including:

- gist perception (Clarke & Mack, 2014),
- perception of ensemble properties (Alvarez, 2011; Haberman & Whitney, 2012), and
- peripheral indeterminacy in vision (Freeman & Simoncelli, 2011).

However, these models rely on a particular computational approximation to Bayes-optimal probabilistic inference, and it is not currently known which probabilistic format the computational processes implemented by the brain have. Hence, future research should especially be targeted at determining which specific type of predictive processing model is implemented by the brain (if any). This will then show which (of the above-mentioned, possible) types of unique constraints on ascriptions of cognitive are provided by PP.

Summary of key points:

1. The paper discusses the relevance of (subpersonal) computational descriptions to (personal) descriptions of beliefs and conscious states.
2. One central result, and the main claim, is that insights can be gained by focusing on the *unique* constraints computational descriptions in PP put on ascriptions of cognitive contents.
3. Unique constraints are provided by, among others, the precision-mediated interplay of estimates in the PP processing hierarchy and (at least in some PP models) by the central importance of summary statistics.
4. These computational posits are relevant to explanations of phenomena like gist and ensemble perception, as well as attention, emotions, or delusions.
5. Future work should especially seek empirical evidence regarding the format of probabilistic inference implemented by the brain (e.g., whether it involves a free-form or a fixed-form code, see Friston, 2009, p. 297-299).

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Structure and Function*, *218*(3), 611–643. doi: 10.1007/s00429-012-0475-5

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. doi: 10.1016/j.tics.2011.01.003

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390-398. doi: 10.1016/j.tics.2012.05.003

Clark, A. (2013). The many faces of precision (replies to commentaries on “whatever next? neural prediction, situated agents, and the future of cognitive science”). *Frontiers in Psychology*, *4*, 270. doi: 10.3389/fp- syg.2013.00270

Clark, A. (2016). Surfing uncertainty: Prediction, action, and the embodied mind. New York: Oxford University Press.

Clarke, J., & Mack, A. (2014). Iconic memory for the gist of natural scenes. *Consciousness and Cognition*, *30*, 256–265. doi: 10.1016/j.concog.2014.09.015

Egan, F. (2014). How to think about mental content. *Philosophical Studies*, *170*(1), 115-135. doi: 10.1007/s11098- 013-0172-0

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*. doi: 10.3389/fnhum.2010.00215

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195-1201. (10.1038/nn.2889) doi: 10.1038/nn.2889

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. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. doi: 10.1038/nrn2787

Friston, K., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological Cybernetics*, *102*(3), 227–260. doi: 10.1007/s00422-010-0364-z

Friston, K., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry*, *1*(2), 148–158. doi: 10.1016/S2215-0366(14)70275-5

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, *193*(2), 559-582. doi:10.1007/s11229-015-0762-9

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness. Searching with Anne Treisman* (pp. 339–349). New York, NY: Oxford University Press.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hohwy, J., Paton, B., & Palmer, C. (2015). Distrusting the present. *Phenomenology and the Cognitive Sciences*, 1–21. doi: 10.1007/s11097-015-9439-6

Madary, M. (2012). How would the world look if it looked as if it were encoded as an intertwined set of probability density distributions? *Frontiers in Psychology*, *3*(419). doi: 10.3389/fpsyg.2012.00419

Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*(9), 451–456. doi: 10.1016/j.tics.2014.05.006

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565–573. doi: 10.1016/j.tics.2013.09.007

Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, *1224*(1), 22–39. doi: 10.1111/j.1749-6632.2011.05965.x

Wiese, W. (2016). What are the contents of representations in predictive processing? *Phenomenology and the cognitive sciences*. doi:10.1007/s11097-016-9472-0

Complementarity Between the Global Workspace Theory and the Sensorimotor Theory

Analyzing workspace theories and sensorimotor theories, Degenaar and Keijzer claim that the two approaches are complementary rather than competitive. They argue that their combination has a higher explanatory power than each theory on its own. The global workspace theory accounts for the difference between conscious and unconscious processes, whereas the sensorimotor theory explains differences between various sensory modularities. Degenaar and Keijzer distinguish three possible ways of combining global workspace and sensorimotor theories: an internal localization scenario, in which experience is localized internally, i.e. in the brain, an external localization scenario, in which experience is localised externally, i.e. outside the brain, and no localisation scenario, in which experience is understood as something that people do rather than as something happening inside their body or as some events generated by their neural system [1].

I focus on two epitomes of these theories: the version of the global workspace theory developed by Baars and his colleagues [2,3,4,5,6,7] and the sensorimotor theory developed primarily by O'Regan and Noë [8,9,10]. Baars’s theory analyses the nervous system as a distributed parallel system, in which many different specialized processes take place. Coalitions of these processes compete for being conscious. The winner of these competitions occupies a so-called global workspace, whose contents are broadcast to other processors and become conscious. The sensorimotor theory, on the other hand, follows the idea of the embodied mind, since the theory stresses an important role of the entire body in which consciousness arises. It opposes a strict division between the brain, which would be supposed to generate mental states, and the rest of the body.

In my analysis, I follow Baars in treating consciousness as a variable that may be predicated on various processes [4]. This naturalist approach determines a specific attitude towards the explanatory gap [11,12].

As noted by Degenaar and Keijzer [1], the strength of the sensorimotor theory lies in its focus on explaining the comparative gap since the theory tries to explain differences between experiences associated with different sensory modalities. Hurley and Noë claim that differences between qualitative character of experience across different sensory modalities (i.e. intermodal differences) can be explained through differences between dynamic patterns of sensorimotor contingencies that are true for particular sensory modalities. This explanation provides an insight into the reason behind different qualitative feels associated with each of the senses. For instance, it is understandable for us on a personal level why a certain set of characteristic sensorimotor contingencies is associated with vision and why a different set of contingencies is associated with hearing [13]. This explains the intermodal comparative gap, i.e. the question why particular neural activity gives rise to a sensation within a particular sensory modality (e.g. vision) rather than within another sensory modality (e.g. hearing).

The global workspace theory focuses on the absolute gap, since it attempts to explain why some neural processes are linked with conscious experience and some others are not [1]. Neural processes compete for entering the global workspace. The winner of the competition broadcast its content through long-range connections to various parts of the central nervous system. This global broadcasting is responsible for conscious experience.

I argue that the global workspace theory and the sensorimotor theory may be combined with predictive processing in its version adopted by Andy Clark [14,15,16,17] under the external localisation scenario. The predictive processing framework helps to explain intricate relations between conscious and unconscious processes. These processes primarily take place within the body, yet are directly related to processes taking place outside the body since the processes of perception and action require constant predictions and estimations of data available in the world. The predictive processing theory unifies exteroceptive, interoceptive and proprioceptive data within a hierarchical structure of multiple processes. It provides a useful link between the global workspace theory, which concentrates on the division into conscious and unconscious processes with a special focus on how consciousness emerges out of dynamic interactions between unconscious processes, and the sensorimotor theory that emphasises human perspective in an active engagement in the world, out of which conscious experience arises. Predictive coding explains processes that underlie conscious and unconscious experience on a more fine-grained level than the sensorimotor theory. In the same vein, conscious processes on the highest level in this multilevel structure accounted for by predictive processing may be identified with the global workspace described by the global workspace theory.

The advantages of combining the global workspace theory with the embodied mind approach may be found in Shanahan's analysis of embedding global workspace architecture within a spatially and temporally located organism. The global workspace architecture leaves unanswered the question why contents that enter the global workspace are characterised by the perspectival unity typical for conscious information. The explanation where this unity comes from may be found once the global workspace architecture "is bound to the spatially confined body". Such an embodied global architecture gives rise to an experiencing subject [18].

I argue that elements of embodied approach may be provided by adding to the global workspace architecture some elements of action-oriented predictive processing and the sensorimotor theory. The sensorimotor theory concentrates on the interaction of the entire organism with the environment. Predictive processing describes events not only in the brain, but also outside the skull since the dorsal horn of the spinal cord is usually indicated as the place where sensory prediction error is generated. Consequently, the holistic approach combining elements of the three theories describes experience as generated by structures outside the brain, so it follows the external localisation specified by Degenaar and Keijzer [1].

In reference to Rao and Ballard's model of predictive coding in the visual cortex, Clark presents a model of an organism organised on various levels within the framework of predictive coding. In this multi-layered model, top-level predictions are about discrete and abstract issues that tend to be more extended in time and diffused in space. Lower-level predictions concern matters that are spatially and temporally continuous, local, and fine-grained. Between the highest and the lowest levels, there is a plethora of many possible intermediate levels [16,17]. I argue that this architecture of multiscale dynamical complexity could be incorporated into the schema in which various coalitions exchange information that sometimes enters the global workspace and becomes conscious. The highest levels, which are in constant communication with lower levels that follow predictive processing, might be equated with the global workspace level. The content of these highest levels constituting the global workspace becomes phenomenally experienced on a personal level and responds to interactions with the environment describable within the sensorimotor theory. Various pathways that carry predictive errors from lower processors and sensory organs may terminate in processors within the cortico-thalamic complex, whose contents compete for entering the highest level - the global workspace - and thereby for becoming conscious.

The cerebral processes within the global workspace are internal processes contributing to consciousness. The interaction of the entire body with the environment described by the sensorimotor theory provides a general description of external processes involved in conscious experience. Both internal and external processes can be experienced at a personal level. The link between external and internal processes on a neuronal level, i.e. at a subpersonal level, is explained within the version of predictive processing developed by Clark.

References:

1. Degenaar, Jan, Keijzer, Fred: "Workspace and Sensorimotor Theories. Complementary Approaches to Experience." *Journal of Consciousness Studies.* 16 (9), 77-102 (2009)

2. Baars, Bernard J.: *A Cognitive Theory of Consciousness*. Cambridge University Press, Cambridge, MA (1988)

3. Baars, Bernard J.: "In the Theatre of Consciousness. Global Workspace Theory, A Rigorous Scientific Theory of Consciousness." *Journal of Consciousness Studies*, 4 (4), 292-309 (1997)

4. Baars, Bernard J.: *In the Theatre of Consciousness. The Workspace of the Mind*. Oxford University Press, New York, NY & Oxford (1997)

5. Baars, Bernard J.: "The Conscious Access Hypothesis: Origins and Recent Evidence." *TRENDS in Cognitive Sciences*. 6.1, 47-52 (2002)

6. Baars, Bernard J.: “The Global Workspace Theory of Consciousness.” In: Max Velmans, Susan Schneider (Eds.) *The Blackwell Companion to Consciousness*, pp.236-246. Blackwell, Oxford (2007)

7. Baars, Bernard J., Franklin, Stan, Ramsoy, Thomas Zoega: “Global Workspace Dynamics: Cortical “Binding and Propagation” Enables Conscious Contents”. *Frontiers in Psychology* 4, 1-22 (2013)

8. O'Regan, Kevin J., Noë, Alva: “A Sensorimotor Account of Vision and Visual Consciousness.” *Behavioral and Brain Sciences*. 24, 939-1031 (2001)

9. Noë, Alva. *Action in Perception.* The MIT Press, Cambridge (MA) and London (2004)

10. O'Regan, Kevin J.: *Why Red Doesn't Sound Like a Bell. Understanding the Feel of Consciousness.* Oxford University Press, New York (2011)

11. Levine, Joseph: "Materialism and Qualia. The Explanatory Gap." In: David J. Chalmers (Ed.). *Philosophy of Mind. Classical and Contemporary Readings*, pp.354-359. Oxford University Press, New York (NY) (1983/2002)

12. Levine, Joseph: “The Explanatory Gap”. In: Brian P. McLaughlin, Ansgar Beckermann, Sven Walter (Eds.) *The Oxford Handbook of Philosophy of Mind*, pp.281-291. Oxford University Press, Oxford (2009)

13. Hurley, Susan, Noë, Alva: "Neural Plasticity and Consciousness". *Biology and Philosophy*. 18, 131-168 (2003)

14. Clark, Andy (2013). “Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science.” *Behavioral and Brain Sciences*. 36, 181-253.

15. Clark, Andy: "Embodied Prediction." In: Metzinger T., Windt, J.M. (Eds.). *Open MIND*: 7(T). MIND Group, Frankfurt am Main (2015)

16. Clark, Andy: "Radical Predictive Processing." *The Southern Journal of Philosophy*. 53, Suppl. 3-27 (2015)

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

18. Shanahan, Murray: "Global Access, Embodiment, and the Conscious Subject." *Journal of Consciousness Studies.* 12, 46-66 (2005)

Symposium at KogWis 2016:

Formal and Cognitive Reasoning

Christoph Beierle[1], Gabriele Kern-Isberner[2], Marco Ragni[3], and Frieder Stolzenburg[4]

[1]Fakultät für Mathematik und Informatik, FernUniversität in Hagen, 58084 Hagen,
[2]Fakultät für Informatik, Technische Universität Dortmund, 44221 Dortmund
[3] RG on the Foundations of AI, Universität Freiburg, 79110 Freiburg
[4]FB Automatisierung und Informatik, Hochschule Harz, 38855 Wernigerode

1 Abstract

Reasoning about (spatial) information is usually pervaded by uncertainty and subject to change. This is not unique to human reasoning, but it also applies to cognitive systems. Thus there is an increasing demand both from psychology and computer science for non-classical reasoning approaches. So far, many advanced and sophisticated approaches of knowledge representation and reasoning have not yet been made accessible for cognitive approaches, and insights gained from cognition are only rarely reflected in formal approaches. Generally, people employ both inductive and deductive reasoning to arrive at beliefs; but the same argument that is inductively strong or powerful may be deductively invalid. Therefore, a wide range of reasoning mechanisms has to be considered. The field of knowledge representation and reasoning offers a rich palette of methods for uncertain reasoning both to describe human reasoning and to model AI approaches. Beyond computational aspects, these methods aim to reflect the rich variety of human reasoning in uncertain and dynamic environments.

The aim of this symposium is to address recent challenges and to present novel approaches to uncertain reasoning in its broad sense, including new insights from cognitive psychology, neuroscience, cognitive computing, and human computation, combining psychological models, uncertain (spatial) reasoning, and computer science. Reflecting this focus, the symposium "Formal and Cognitive Reasoning" at KogWis 2016 is organized jointly by the GI special interest groups *Wissensrepräsentation und Schlie"sen* and *Kognition.*

2 List of Speakers

– Francois Bry, LMU München
  Human Computation: Combining the Computational Power of Machine with Cognitive Skills of Humans

Artificial Intelligence is a buzz word of the first two decades of the 21st century. An interesting aspect of many, if not all, of the much celebrated artificial intelligences is their building upon human cognitive skills. Human Computation is a name given to an approach consisting in engaging humans to contribute to software systems. This presentation will report on Human Computation applications and stress essential aspects of successful human computation systems.

– Christian Freksa, Universität Bremen
  Affordance and Constraints as complementary notions in problem solving

  In artificial intelligence (AI) we employ constraint-based reasoning for problem solving. When doing so, we essentially conceptualize a large space of theoretically conceivable values that we restrict to a subset of values that are legal in our specific problem domain. This is an extremely powerful approach for using general-purpose computers or general-purpose programming paradigms to adapt them to the specific requirements of a given problem.
  In psychology, Gibson (1979) introduced the notion of affordances that has been used in the design of human-computer interaction (Norman 1980). An affordance can be seen as a specific spatial config-uration that provides a specific opportunity by permitting (or afford-ing) a specific action.
  In my contribution I want to discuss, (1) to what extent the notions of constraint and affordance can be viewed as complementary; (2) whether they equally apply to physical and computational systems; (3) how they apply to problem solving in humans and machines; (4) how they can be employed in interaction; and (5) what are their relative merits.

– Markus Knauff, Universität Gießen
  New Frameworks of Rationality

  Rationality is a key concept in psychology and philosophy. However, for a long time, a divide and conquer approach between psychology and philosophy has prevented both disciplines from taking into account one anothers progress . Philosophys miss ion has been to characterize what it means to be rational, and to put forth general principles, formal theories, and axioms defining rationality. In contrast, psychologys miss ion has been to empirically investigate to what extent peoples cognition and behavior conforms to those norms of rationality, and if deviations occur, explain why they occur. In my talk, I will give an overview and present some highlights from the DFG-funded Priority Programs New Frameworks of Rationality (SPP1516), in which psychologists and philosophers (and AI researchers) work together to overcome the previous division of labor and to develop and test shared new concepts of human rationality. Our findings show that these new concepts of rationality challenge the conjecture of human beings as intrinsically irrational or illogical. People are often smarter than we think.

3 Symposium Organizers

- Christoph Beierle, FernUniversität in Hagen
- Gabriele Kern-Isberner, Technische Universität Dortmund
- Marco Ragni, Albert-Ludwigs-Universität Freiburg
- Frieder Stolzenburg, Hochschule Harz

Effects of ageing on landmark recognition

Ramona Grzeschik [1,2], Ruth Conroy Dalton[3], Anthea Innes[4], Jan Wiener[1,2]

1 Department of Psychology, Bournemouth University, Bournemouth, UK
2 Bournemouth University Dementia Institute, Bournemouth, UK
3 Department of Architecture & Built Environment, Northumbria University, Newcastle, UK
4 Department of Applied Social Science, University of Stirling, Stirling, UK

rgrzeschik@bournemouth.ac.uk

Landmark recognition is essential for route learning. However, not every object is a useful landmark, be it due to repetitive occurrence or visual appearance. Salient objects (e.g. fire extinguishers) catching attention might be easier to memorize than non-salient objects, but if they appear more than once along a route, they can't be used as reliable landmarks. Cognitive ageing affects executive functions and control of attention which could impact on choosing relevant, but non-salient landmarks and therefore route memory [1, 2]. The aim of the presented study is to investigate how cognitive ageing affects people's ability to select unique objects as landmarks for place and route learning and how deficits in landmark selection might affect the navigation skills of older participants. To do so, we created two kinds of routes through a virtual care home: simple routes and complex routes, each comprising four intersections each with two objects. For simple routes, the unique landmarks (objects that occurred only once along the route) were also salient. For the complex routes, in contrast, the salient objects occurred twice on the route (=non-unique) and the non-salient objects were unique. We recorded route learning performance and gaze-behaviour from young and old (65+) participants. The gaze data revealed that older participants attended less to the unique landmarks on the complex routes than on simple routes, while younger participants primarily attended the unique landmarks, regardless their saliency. This effect was also reflected in the performance data: young participants did not show any performance difference in learning simple and complex routes, whereas the older participants performed better on the simple routes than on the complex routes. These results suggest that cognitive ageing affects the control of visual attention which, in turn, contributes to age-related deficits in route learning performance.

References

1. Verhaeghen P , Cerella J. (2002). Aging, executive control, and attention: a review of meta-analyses. *Neuroscience and Biobehavioral Reviews, 26*, 849-857
2. Zhong J.Y., Moffat S.D. (2016). Age-Related Differences in Associative Learning of Landmarks and Heading Directions in a Virtual Navigation Task. *Frontiers in Aging Neuroscience, 8,* Article 122

The influence of extra-linguistic information on spatial language

Michele Burigo
CITEC, University of Bielefeld, Bielefeld, Germany

Abstract. The apprehension of a spatial description such "The vase is *above* the book" requires extra-linguistic knowledge, e.g. functional relation and inferences people can draw from the geometry of the objects. If the role of the potential interaction between the objects received much attention, the information derived from the objects' geometry has been often neglected. Here I present a number of experimental investigations showing that the geometric information extracted from the located object (i.e., the book) affects how people understand and produce spatial descriptions.

Keywords: spatial language, spatial prepositions, extra-linguistic inferences, located object, reference object, object geometry.

1 Spatial language and objects' geometry

Language is the principal tool human beings use for communication. Spatial language is the branch of language used to communicate information about space and locations. "The vase is *above* the book", "The car is *in front of* the house", "The sofa is *on the left of* the television" are common sentences people use to describe objects' location. These examples focus is on the location of the located object (the vase, the car and the sofa, in the examples above), which is the subject of the descriptions. The other object, the so-called reference object (the book, the house and the television), is used as a spatial reference to narrow the search for the located object.

In order to understand spatial descriptions, in addition to the linguistic contents, people also take into account object knowledge and situational information [1, 2, 3]. Object knowledge includes the geometric properties extracted from the reference object and researches have shown that its rotation is critical for the interpretation of spatial relations based on the intrinsic reference frame [4, 5, 6, 7]. On the other hand the geometric information extracted from the located object has been often neglected. However there is evidence that also the geometric information derived from the located object might be relevant for the spatial language processing [3], [8, 9]. Following this evidence, we conducted a number of studies investigating whether the rotation of the LO affects spatial language apprehension.

1.1 The role of the geometry of the located object for spatial language

In the first series of studies we asked participants to place the located object in a target location defined by a spatial description such as “The pumpkin is *above* the strawberry” (see Fig. 1). Another group of participants were asked to produce a spatial description of those scenes, while the time necessary to utter the spatial description was recorded. The results showed that scenes where the located object was presented with a rotation in contrast with the one used for the reference object (“non-canonical LO” in Fig. 1) received lower acceptability ratings and took longer to be described [10].

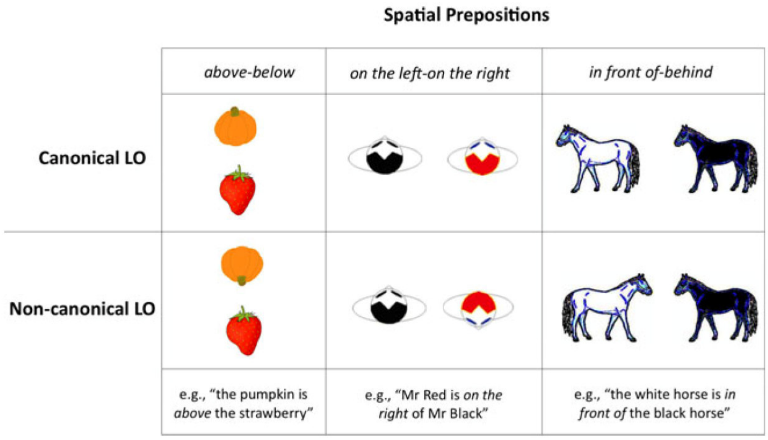


Fig. 1. Examples of scenes where the LO presents the same rotation as the RO (canonical LO) and scenes where the LO has a different rotation (non-canonical LO) across three sets of spatial terms. Reprinted with permission from “Object Orientation Affects Spatial Language Comprehension” (p. 1475), by M. Burigo & S. Sacchi, 2013, Cognitive Science, 37. Copyright 2013 by Cognitive Science Society.

In a following study we further investigated the role of the geometry of the located object by examining whether longer placing time and production latencies were related to the computation of an inference that has been shown to be relevant in language comprehension: *converseness*. This logic states that if “A is *above* B” is true, then “B is *below* A” must also be true [9]. Under normal circumstances converseness holds, but when the LO is rotated (non-canonical LO, Fig. 1) it does not. In fact if the description “The pumpkin is *above* the strawberry” is acceptable, the converse description “The strawberry is *below* the pumpkin” is, at least from the intrinsic perspective, not acceptable. An acceptability rating task showed that descriptions presented together with a scene where converseness held received higher ratings compare to scenes where the logic of converseness did not hold [11].

Finally we looked more in detail at the geometric information extracted from the LO that might have triggered the conflict. To do so we asked participants to place the located object in a target location while the information concerning the orientation and the direction of the located object were manipulated. Preliminary results seem to indicate that is only the direction to be important, despite the fact that orientation and direction may be organized in a hierarchical manner [12].

In summary we showed that the understanding of a spatial description such as “A is *above* B” is modulated by information that goes beyond language, such as the orientation of the LO.

References

1. Carlson-Radvansky, L. A., & Radvansky, G. A. (1996). The influence of functional relations on spatial term selection. *Psychological Science*, 7, 56–60.
2. Coventry, K., & Garrod, S. (2004). *Saying, seeing and acting*. Hove, England: Psychology Press, Taylor & Francis Group.
3. Coventry, K., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of over, under, above, and below. *Journal of Memory and Language*, 44, 376–398.
4. Carlson, L. A. (1999). Selecting a reference frame. *Spatial Cognition and Computation*, 1, 365–379.
5. Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment. *Journal of Memory and Language*, 33, 646–671.
6. Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction. *Journal of Memory and Language*, 37, 411–437.
7. Carlson, L. A., & Van Deman, S. R. (2008). Inhibition within a reference frame during the interpretation of spatial language. *Cognition*, 106, 384–407.
8. Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition*, 31, 45–60.
9. Levelt, W. J. M. (1984). Some perceptual limitation on talking about space. In A. J. van Doorn, W. A. van der Grind, & J. J. Koenderik (Eds.), *Limits on perception* (pp. 323–358). Utrecht, the Netherlands: VNU Science Press.
10. Burigo & Sacchi, 2013. Object Orientation Affects Spatial Language Comprehension, *Cognitive Science*, 37 (2013) 1471–1492.
11. Burigo, Coventry, Cangelosi & Lynott, 2016; Michele Burigo, Kenny R. Coventry, Angelo Cangelosi & Dermot Lynott (2016): Spatial language and converseness, *The Quarterly Journal of Experimental Psychology*, DOI: 10.1080/17470218.2015.1124894.
12. Burigo & Schultheis (in preparation). The effects of direction and orientation of the located object on spatial language comprehension.

Adjusting our view on perspective taking: Scalable representation structures and reference frames

Holger Schultheis

Cognitive Systems, Bremen University, Germany

schulth@sfbtr8.uni-bremen.de

The ability to judge spatial arrangements from vantage points that differ from one's actual bodily location and orientation has previously been assumed to rely on the ability to mentally transform (translate, rotate) body location and orientation. More recent studies have cast doubt on such a mental transformation account of perspective taking. I will present a computational model that explains perspective taking in terms of a joint influence of interference on the cognitive and the motor level. I will also discuss possible representational bases giving rise to these interference processes. These considerations indicate commonalities between perspective taking and other spatial abilities that so far have only been considered in isolation of each other.

The decomposition of navigation behavior into simple tasks

Jascha Gruebel, Tyler Thrash, Victor Schinazi, and Christoph Hoelscher

Chair of Cognitive Science, ETH Zurich, Switzerland
jgruebel@student.ethz.ch

Abstract. Research in spatial cognition often defines participants' tasks in terms of underlying cognitive components. However, the extent these different tasks combine to explain relatively complex navigation behavior is largely unknown. We attempted to predict participants' navigation performance in a virtual environment using eight simple tasks that were defined according to four common distinctions. Each task emphasized egocentric or allocentric reference frames, perceptual or cognitive processes, directions or distances, and static or dynamic stimuli. For example, an allocentric, cognitive, distance task with static stimuli required participants to move towards a target that was marked on a map from memory. In order to assess the validity of this theoretically informed task structure, we compared its task associativity with similar random models. In contrast to our expectations, our original model performed only average compared to the random models. We also determined the optimal model in terms of task associativity by performing Principal Component Analysis on participants' performances for these tasks. This "optimal" model differed from our original model in that the primary driver of spatial performance seemed to be the abilities to read maps and to consider dynamic stimuli. These results may be explained by a discrepancy between the way in which the tasks were operationalized and the structure represented by the four distinctions. Consequently, either the task design or the distinctions from the literature need to be improved in order to generalize to more complex navigation behavior.

The integration of room views

Chantal Horeis[1], Celia Foster[1], Katsumi Watanabe[2], Heinrich H. Bülthoff[1], Tobias Meilinger[1,2]

[1] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
{chantal.horeis, celia.foster, heinrich.buelthoff, tobias.meilinger}@tuebingen.mpg.de
[2] The University of Tokyo, Tokyo, Japan
{kw, tm}@rcast.u-tokyo.ac.jp

Rooms cannot be experienced within a single view as humans cannot look backwards. Nevertheless, humans are able to form an understanding of the whole room. We were interested how and under which conditions integrated room representations are formed based on which long-term memory structure. Participants experienced views of a rectangular virtual room from its center through a head mounted display. Afterwards, they saw a room view and indicated the direction of a non-visible room object using the arrow keys of a keyboard. Participants responded quicker for the first view encountered than for later experienced room views. This pattern did not change when participants rotated physically during learning or only visually. These results indicate that participants did not update experienced room views during learning to memorize integrated room information and are therefore inconsistent with integrating cognitive map parts via path integration [1]. The results are consistent with memorizing separate room views and the transitions between them [2] as well as with memorizing an integrated room memory in a reference frame oriented along the first experienced room view [3]. Our data cannot clearly separate between the two possibilities suggesting that both strategies might have taken place to some degree. The model best fitting with the data suggests that integrating within a single reference frame most often occurred when participants could look around in a self-determined sequence as long as they wanted in continuously changing perspectives. Contrary, when the sequence of views was pre-determined participants most often relied on a sequence of stored views. In sum, results indicate, firstly, that humans do not necessarily integrate experienced room views during learning, even they know that they have to act on an integrated room representation afterwards. Secondly, the first experienced room view acts as an anchor later experienced views are related towards. Thirdly, spatial long-term memory formation seems ra-

adfa, p. 1, 2011.

ther independent from updating the spatial surrounding in working memory during learning.

References

1. Wang, R. F. (2016). Building a cognitive map by assembling multiple path integration systems. *Psychonomic Bulletin & Review, 23,* 692–702.

2. Mallot, H. A. (1999). Spatial cognition: Behavioral competences, neural mechanisms, and evolutionary scaling. *Kognitionswissenschaft, 8,* 40-48.

3. Kelly, J. W., & McNamara, T. P. (2010). Reference frames during the acquisition and development of spatial memories. *Cognition, 116,* 409–420.

What is Orientation?

Jakub Krukar and Angela Schwering

Institute for Geoinformatics, University of Muenster,
Heisenbergstr. 2, 48149 Muenster, Germany;
krukar@uni-muenster.de

Keywords: orientation, wayfinding, computer-assisted navigation

The dominant approach for computer-assisted navigation is realised by a turn-by-turn set of instructions communicated briefly before the decision point of the navigator. This technology has been shown to have no positive impact on users' survey knowledge [5], even after prolonged periods of active usage. One suggested reason for such an effect is its incompatibility with natural means of structuring and communicating spatial knowledge exhibited by human navigators [6]. While expressing wayfinding instructions, people use a wide variety of landmark types, simplified qualitative relations between route elements, refer to distant but salient spatial features as well as to a broad range of other 'auxiliary' information. These additional units of information are often not required to give the 'minimal correct' route description, and yet, navigators show consistency in the selection of elements deserving such a mention.

The recently commenced project 'WayTO: Wayfinding Through Orientation' aims to prototype and test a navigational support system whichadditionally to enabling an effective travelwill positively contribute to spatial orientation of its user in their broader urban environment. This will be realised by enriching navigational information with elements consistently used by human navigators for structuring and communicating spatial knowledge.

Traditionally, researchers have considered 'orientation' as synonymous to 'human performance in tasks' measuring concepts as diverse as spatial knowledge, spatial abilities and wayfinding performance. As a result, the existing literature uses the term 'orientation' inconsistently, sometime meaning a vague mixture of the above components contributing to one's understanding of their own location, and other times as a synonym of very narrowly defined notions such as 'survey knowledge'.

While we believe that survey knowledge (or spatial knowledge in general) is an important component of 'orientation', the terms are here distinguished. This is guided by the fact that complete spatial knowledge does not guarantee a perfect and continuously correct understanding of one's own location in the broader urban context. Conversely, a 'good enough' orientation is possible (and often exhibited) by individuals having a very limited knowledge of the environment. In the narrowest existing definitions, considering 'orientation' synonymous to 'survey knowledge' would indicate that a large proportion of population is as bad in remaining oriented as is their survey knowledge of the considered environment (and consequently their performance in survey knowledge-centred tasks employed in wayfinding studies). This seems to be untrue - as both research evidence and practical observations suggest, human navigators can flexibly adopt a broad range of strategies, combine unstructured, biased, hierarchical spatial knowledge of uneven quality or certainty and infer unknown spatial properties in order to reach their destinations.

This is possible as an imperfect spatial knowledge of one's own surrounding is not a 'mental map with blacked-out regions' but a mixture of graph-based representation of relations between regions and features, partial and intuitive knowledge of qualitative spatial relations, knowledge of hierarchical structure of objects in regions, and metric information at the finer level of detail [3, 4, 7]. And yet, despite a significant advancement in our understanding of the structure of spatial mental representations, wayfinding studies and technological evaluations remain dominated by measures relying on route- or survey-knowledge: the implication often being, that a more complete 'cognitive map' is a sign of better orientation. Such an approach does not consider human ability to infer some spatial properties and remains ignorant of the non-uniform importance of the precision and correctness of those properties for the success of natural navigation.

We therefore propose to consider orientation as a dynamic process of deriving one's position in space with regard to known environmental information at a scale (or subset of scales) relevant to the current goal. This implies that such a representation can be correct at some levels of conceptualising space while remaining imprecise (or even incorrect) at others. For instance, a navigator might know that they are located at position A, south from the city centre, and that the destination point B is somewhere inside a region located north-west of the centre. They might be aware of the fact that arriving there from the current position would require following a path along a park and crossing a river. These are all correct informational units sufficient for a high degree of orientation - probably sufficient to complete a navigational task and to update one's location along the route. And yet, externalisation of this representation can result in a broad range of possible results due to the need to infer unknown spatial information (e.g. the angular direction from A to B or the exact location of B inside its region). Figure 1 demonstrates some sample possibilities.

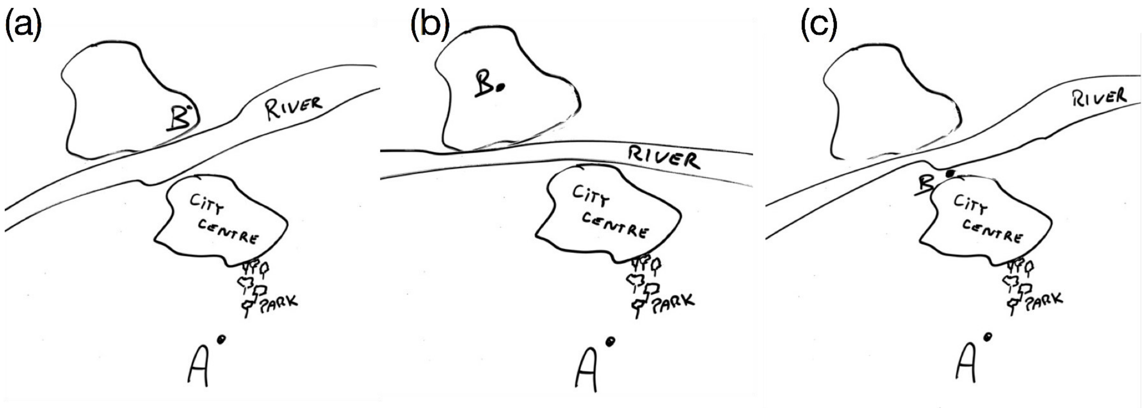


Fig. 1. (a) Correct survey representation of the environment. (b) Potential externalisation variant based on participant's imperfect knowledge assuming the navigator only knows that they are located at position A, south from the city centre, and that the destination point B is 'somewhere inside' a region located north-west of the centre, as well as that arriving there would require following a path along a park and crossing a river. Estimated position of point B in (b) might be biased towards the centre of the region and river might be adjusted to the orthogonal axis. This externalisation would yield relatively large errors in a Pointing Task and Bidimensional Regression Analysis of sketch maps. Despite being incorrect (or imprecise), it should suffice to orient in the environment. (c) A small metric, but large qualitative error, inconsistent with participant's knowledge. This would yield good survey-knowledge results if measured traditionally.

Note that as traditional measures used for establishing wayfinding performance are dominated by the 'map in the head' metaphor, they could yield drastically different results. For instance, poor pointing performance or a low bidimensional regression score of a drawn sketch map can be observed even when the actual knowledge of the required spatial relations is matching the factual environmental configuration at a degree sufficient for a navigational task. The opposite is also possible: a high bidimensional regression score or pointing performance can be observed, while participant has an incorrect representation of key spatial relations between important structural elements of the environment; for instance, believing that the destination is at the incorrect side of the river would correspond to a small metric/angular error in traditional tasks but jeopardise one's navigational possibilities in the region.

Despite the vagueness resulting from imperfect and uneven knowledge of different spatial properties, humans are relatively efficient at using such sparse and unevenly distributed knowledge to navigate. One strategy demonstrating that being the on-line correction of wrong assumptions based on the newly gained visuo-spatial information during walking [2].

Considering orientation as a dynamic process and not as a stable, constantly updated representation on a metric 'map in the head', implies that orientation can vary on a goal-by-goal (or rather 'extraction-by-extraction') basis. Asking a participant to point to an element X can prompt different orientation than pointing to element Y as element Y might be strongly associated with previously unneeded features or relations (e.g. at a larger scale). The process of inferring unknown spatial properties can be thus seen as not less important than the process of externalising known spatial elements and relations [1].

In this view, an orientation-supportive navigational assistance system can contribute to its user's orientation by:

a) correcting incorrect (biased) assumptions about a subset of those spatial properties which might be most broadly applied to other potential cases of deriving orientation (e.g. global and regional landmarks, large structural features);

b) linking unknown spatial knowledge to known spatial features;

c) highlighting information about yet-unknown structural regularities and hierarchies assisting to organise the newly acquired knowledge in a manner in-line with the ways such knowledge is organised naturally (for instance in a hierarchical way dominated by salient features, and alignments to regular geometrical shapes).

In the talk we will present the goal of the recently commenced 'Wayfinding Through Orientation' project, preliminary results supporting some of the above claims, as well as theoretical and methodological considerations for the way forward.

Acknowledgement

This research is supported by the 'Wayfinding Through Orientation' ERC StRG Grant Agreement No 637645.

Bibliography

[1] Appleyard, D.: Styles and methods of structuring a city. Environment and Behavior 2(1), 100–117 (1970)

[2] Hoelscher, C., Tenbrink, T., Wiener, J.M.: Would you follow your own route description? Cognitive strategies in urban route planning. Cognition 121(2), 228–247 (2011)

[3] Ishikawa, T., Montello, D.R.: Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. Cognitive psychology 52(2), 93–129 (2006)

[4] Meilinger, T.: The Network of Reference Frames Theory: A Synthesis of Graphs and Cognitive Maps. In: Spatial Cognition VI. Learning, Reasoning, and Talking about Space, pp. 344–360. Springer Berlin Heidelberg, Berlin, Heidelberg (2008), http://link.springer.com/10.1007/978-3-540-87601-4_25

[5] Münzer, S., Zimmer, H.D., Baus, J.: Navigation assistance: A trade-off between wayfinding support and configural learning support. Journal of experimental psychology: applied 18(1), 18 (2012)

[6] Schwering, A., Li, R., Anacta, V.J.A.: Orientation information in different forms of route instructions. In: Short Paper Proceedings of the 16th AGILE Conference on Geographic Information Science, Leuven, Belgium (2013)

[7] Tversky, B.: Cognitive maps, cognitive collages, and spatial mental models, pp. 14–24. Springer Berlin Heidelberg, Berlin, Heidelberg (1993), http://dx.doi.org/10.1007/3-540-57207-4_2

Digit Position and Force Synergies during Unconstrained Grasping

Abdeldjallil Naceri[1,2] , Alessandro Moscatelli[1,], Robert Haschke[2], Marco Santello[3], and Marc Ernst[1,2]

[1]Department of Cognitive Neuroscience, Bielefeld University, 33615 Bielefeld
[2]Center of Excellence Cognitive Interaction Technology (CITEC), Bielefeld University, 33619 Bielefeld
[3]School of Biological and Health Systems Engineering, Arizona State University Tempe, Arizona 85287

Grasping is a complex motor task which requires a fine control of the multiple degrees of freedom of the hand, in both the position and the force domain. In this work, we investigated the digit force controller type in the human hand while grasping and holding a perturbed object. We observed a substantial variability between participants in the hand posture. Instead, digit placement was rather stereotyped for repeated grasps of the same participant. The magnitude of digit normal force corrective responses was controlled in a feed-forward manner within single perturbations. Hand net torque was optimized across perturbations by increasing the grip force in the perturbations following the first one, which indicates that it was controlled in a closed-loop fashion. The synchronous increase of digit normal forces was achieved by all digits regardless of the external perturbation and the number of digits involved in the task. This suggests that the CNS adopted a stiffening strategy that compensates for the task uncertainty associated with unpredictable external perturbations and with the variability in required grip force due to the natural trial-to-trial variability.

Anticipating Object Interaction with the Eyes and with the Hands: Perceptual and Planning Aspects

Anna Belardinelli[1], Stefanie Blumenschein[2], and Martin V. Butz[3]

[1]Department of Cognitive Neuroscience, Bielefeld University, 33615 Bielefeld
[2]Center of Excellence Cognitive Interaction Technology (CITEC), Bielefeld University, 33619 Bielefeld
[3]School of Biological and Health Systems Engineering, Arizona State University Tempe, Arizona 85287

During object manipulation our eye movements are directed to the target of the interaction and to the relevant sites where this takes place. Specifically, when reaching to grasp an object, the overt attention focus is mostly anchored at the location where the index finger will be placed [1], to help monitoring and directing the effector and anticipating the first contact with the object. Thus, the oculomotor behavior anticipates the upcoming sensorimotor contingencies signaling the satisfied precondition for the next planned piece of action.

On the other hand on the motor side, studies on the End-state comfort effect [2] have shown how the object interaction to be carried out with the grasped object also influences the reaching and grasping phase. Indeed, an uncomfortable grasp is planned in advance and executed, whenever the final posture turns out to be more kinematically convenient for the following interaction.

All these effects pertain to object-related actions we execute in our peripersonal space. In this space multisensory processing and, specifically, visuo-tactile interactions play a fundamental role, acting as an anticipative sensory-motor interface. These interactions have been studied by means of the cross-modal congruency task, where a visual and a tactile stimulus are simultaneously congruently/incongruently delivered, the former at the location where either the thumb or the index finger will make contact during a grasp, the latter either on the thumb or the index finger of the grasping hand. Incongruent conditions are characterized by longer reaction times when answering which finger was haptically stimulated. [3] postulated that this congruency effect is the sign of a remapping of multisensory peripersonal space coding, online modulated by the difficulty of the performed action (grasping vs. pointing) and by the SOA of the stimulation.

In this study we tested whether not just the difficulty of the action but the more or less complicated way this is planned also modulates the remapping.We asked participants to pantomime a grasping and placing task of a bottle in front of a touch screen and to concurrently attend to the perceptual task of identifying which finger was stimulated by a vibro-tactile device. The tactile stimulus was accompanied by the appearance of small dot (visual stimulus) on either the right or the left side of the bottle, which could be presented either upright or upside down. This latter manipulation would induce a thumb-up or thumb-down grasp, respectively. We hypothesized that not only the visuo-tactile congruency would have an effect but also the stimulated side/finger. The visual stimulus could indeed be on the side where the index finger would land (right in the upright, left in the upside condition) or the congruency effect could be favored by attention on the index finger. Response times, eye-tracking and motion-tracking data were collected.

Preliminary results show effects of congruency, with shorter response times for congruent conditions. A stronger cross-modal congruency effect was also found when the stimulation was delivered during the reaching phase as compared to before or at movement onset. Movement times were longer in the inverted condition and for earlier SOAs. Eyetracking results are currently analyzed. These results seem to support the hypothesis that our mind remaps the peripersonal space around the hand in an anticipatory manner with respect to the planned, upcoming object grasp (including an end-state oriented grasp when the object is presented standing on its top).

References

1. Cavina-Pratesi, C., Hesse, C.: Why Do the Eyes Prefer the Index Finger? Simultaneous Recording of Eye and Hand Movements during Precision Grasping. J. of Vision 13, 1-15 (2013)
2. Rosenbaum, D.A., Marchak, F., Barnes, H.J., Vaughan, J., Slotta, J.D., Jorgensen, M.J.: Constraints for Action Selection: Overhand versus Underhand Grip. In: Jeannerod, M. (eds.) Attention and performance XIII, pp. 321-342. Lawrence Arlbaum Associates, Hillsdale (1990)
3. Brozzoli C., Cardinali L., Pavani F., Farne A.: Action-Specific Remapping of Peripersonal Space. Neuropsychologia 48, 796–802 (2010)

The Neurophysiological Interaction between Working Memory and Grasping Movements

Rumeysa Gunduz-Can[1,2], Thomas Schack[1,2,3], and Dirk Koester[1,2]

[1]Center of Excellence – Cognitive Interaction Technology (CITEC), Bielefeld University;
[2] Faculty of Psychology and Sport Science, Bielefeld University;
[3]Research Institute for Cognition and Robotics, Bielefeld University

For human beings, manual actions have been the key for interacting with the physical world. We reach for objects at different distances, manipulate objects depending on our goals, or grasp objects with different shapes or weight. In natural environments, such actions are rarely performed in isolation. Rather, they are performed under complex, dynamic situations that require higher cognitive operations such as grasping a coffee cup while talking to a friend. Moreover, dynamic situations such as a change in the target's shape or orientation may demand re-planning of the ongoing action, i.e., adapting the current plan to the changing action demands. So far, it has not been fully understood how the human brain accomplish flexible, yet precise manual actions in complex, dynamic situations. Therefore, we aim to further our understanding of the neuro-cognitive mechanisms behind manual actions and action flexibility. Particularly, we are interested in the neurophysiological interaction between working memory (WM) and grasping movements which are the most complex and cognitively organized manual actions.

Here, we present two studies in which we focused on the different movement components (planning, execution, re-planning), different WM domains (verbal, visuospatial) and processes (encoding, maintenance, retrieval). For both studies, we implemented a dual-task scenario which requires simultaneous performance of a WM task and a grasp-to-place task (See Fig. 1). Participants, first, grasped a sphere and planned a placement movement toward either the left or right motor target. Then, they encoded the visually presented WM stimuli, i.e., either letters for the verbal WM task or symbols within a matrix for the visuospatial WM task. Subsequently, participants performed the placement movement while maintaining the information in WM. Finally, they reported the information after finishing the movement task.

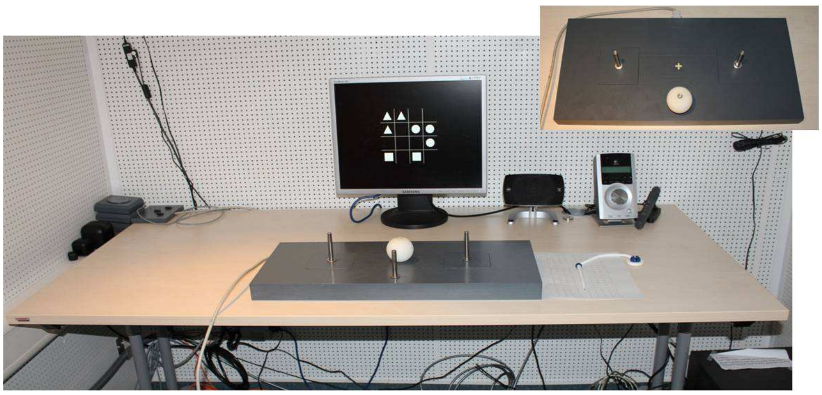


Fig. 1. The experimental set-up and task board for the grasp-to-place task (top view). The task board included two sticks (motor targets) and a sphere which is placed onto a start position as shown in the picture.

In the first study, we focused on the neuro-cognitive costs of performing a pre-planned grasping movement for encoding and retrieval of WM. The study rested on 2 (WM domain: Verbal vs. visuospatial) x 2 (Task condition: Single vs. Dual) within-subject design. That is, a baseline single-task condition (the verbal and visuospatial WM task) was compared with a dual-task condition (simultaneous performance of the WM and grasp-to-place task). Thirty participants were tested.

The ERP results showed neurophysiological interference costs of grasping movements only for the encoding process of WM. In the single-task block, the ERP analyses showed a three- way interaction of WM domain, hemisphere and anterior-posterior orientation of ROI between 200-500 ms, $F(1, 24) = 16.18$, $p < .001$, $\eta p = .413$. Following post-hoc analyses indicated the differential cortical activity for verbal and visuospatial WM tasks at bilateral anterior and right posterior recording sites. In the dual-task block, the ERP analyses showed a two-way interaction of WM domain and anterior-posterior orientation of ROI between 200-400 ms, F (1, 24) = 13.06, p =.001, ηp = .352. Following post-hoc analyses indicated the differential cortical activity for the WM tasks at bilateral posterior recording sites (See Fig. 2). Moreover, the single-task ERP analyses add to the converging evidence for the distinct neural sources for encoding and retrieval processes of verbal and visuospatial WM domains. For the verbal domain, the expected anterior negativity was found only during the encoding process. For the visuospatial domain, the expected posterior negativity was shown during both the encoding and retrieval processes [1], [2]. Behavioral data, in accordance with Spiegel, Koester, Schack [3], showed domain-specific movement costs of the grasp-to-place movement for the visuospatial domain. That is, performing additional pre-planned grasping movement decreased memory performance for the visuospatial task, but not for the verbal task.

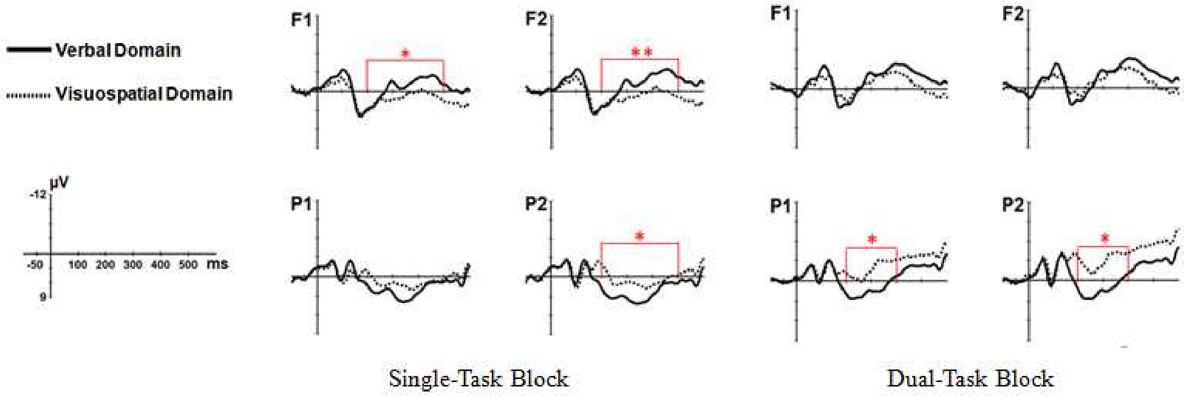


Fig. 2. Encoding process ERPs. One example electrode from each region of interest (ROI). In the single task block, there were bilateral anterior and right posterior effects. In the dual task condition, there was a bilateral posterior effect. * p <.05, ** p <.001 for paired sample t-tests

Generally, this study provides an evidence for process-specific grasping interactions with WM during the encoding process. This study, therefore, provides an initial neurophysiological characterization of functional interaction between WM and grasping movements in a complex dual-task setting.

In the second study, we focused on the neuro-cognitive mechanisms of grasping movement flexibility in terms of movement re-planning. Particularly, we focused on the neuro-cognitive costs of implementing a new movement plan during an ongoing grasping for encoding, maintenance and retrieval of WM.

Dual-task scenario was only different from the first study in terms of re-planning requirements. For the 30% of trials, participants had to re-plan the movement direction after encoding the WM stimuli. Therefore, the second study rested on a 2 (WM domain: Verbal and visuospatial) x 2 (Movement planning: Pre-planned and re-planned) within subject design. 35 participants were tested.

Preliminary ERP results showed the main effect of planning condition during the maintenance process, F (1, 30) = 8.057, p =.008, ηp = .212. That is, the ERPs for both the verbal and visuospatial domains changed when there was a need for movement re-planning (See Fig. 3). There was no interaction for encoding and retrieval processes. Behavioral results also supported to the ERP findings showing a lower memory performance for both WM tasks in the re-planned condition compared to the pre-planned condition. We interpret these findings as indicating the neurophysiological interference costs of movement re-planning for the maintenance process.

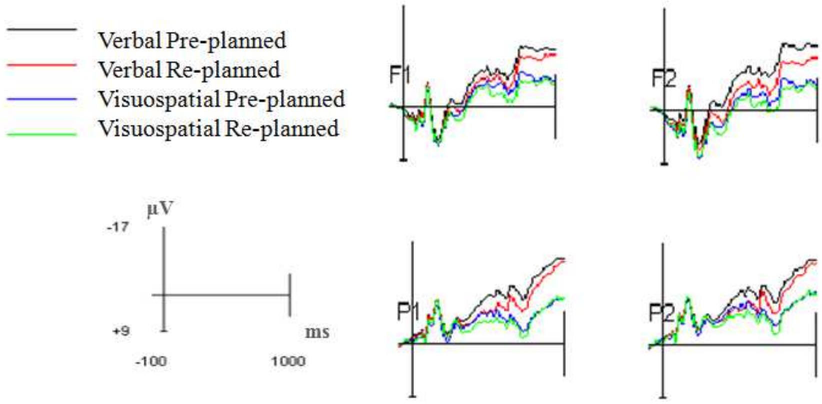


Fig. 3. Maintenance process ERPs. One example electrode from each ROI. Re-planning effect is seen for both WM tasks in each ROI.

With these two studies, we reported reliable ERPs in a complex experimental setting including overt movement execution. In this sense, our studies support and extend the studies analyzing ERPs in mere grasping tasks [4], [5]. More importantly, we showed that movement (re)planning and execution recruit different WM capacities. That is, performing a pre-planned grasping movement mainly interferes with WM during the encoding process, while re-planning the movement, which is more cognitively demanding, interferes with visuospatial and verbal domains during the maintenance process. More generally, these two studies provide an empirical basis for further neurophysiological investigations of neuro-cognitive interactions between WM and (manual) action control.

References

1. Bosch, V., Mecklinger, A., Friederici, A. D.: Slow Cortical Potentials during Retention of Object, Spatial, and Verbal Information. Cognitive Brain Research 10, 219-237 (2001)

2. Pinal, D., Zurrón, M., Díaz, F.: Effects of Load and Maintenance Duration on the Time Course of Information Encoding and Retrieval in Working Memory: From Perceptual Analysis to Post-Categorization Processes. Frontiers in Human Neuroscience 8, 1-15 (2014)

3. Spiegel, M. A., Koester, D., Schack, T.: The Functional Role of Working Memory in the (Re-) Planning and Execution of Grasping Movements. J. of Experimental Psychology: Human Perception and Performance 39, 1326-1339 (2013)

4. Westerholz, J., Schack, T., Koester, D.: Event-Related Brain Potentials for Goal-Related Power Grips. PLoS ONE 8, e68501 (2013)

5. Westerholz, J., Schack, T., Schütz, C., Koester, D.: Habitual vs Non-Habitual Manual Actions: An ERP Study on Overt Movement Execution. PLoS ONE 9, e93116 (2014)

Distinct Effects of Visuomotor Priming on Action Preparation and Motor Programming

Christian Seegelke[1,2], Iris Güldenpenning[3], Julian Dettling[1], and Thomas Schack[1,2,4]

[1]Neurocognition and Action Research Group, Faculty of Psychology and Sport Sciences, Bielefeld University, Bielefeld, Germany
christian.seegelke@uni-bielefeld.de
[2]Center of Excellence Cognitive Interaction Technology (CITEC), Bielefeld, Germany
[3]Sport Psychology Unit, Department of Sports and Health, Faculty of Science, University of Paderborn, Paderborn, Germany
[4]Research Institute for Cognition and Robotics (Cor-Lab), Bielefeld, Germany

Keywords: grasping • visuomotor priming • action • perception • motor planning

The prominent "two visual systems" hypothesis proposes a functional distinction between vision for perception (ventral stream) and vision for action (dorsal stream). According to this model, visually guided actions are computed in real-time and hence, should show no influence from previous experience. In the present study, we examined whether the planning and execution of a reach-to-grasp movement is susceptible to visuomotor priming, and whether the strength of the priming effect depends on the availability of visual feedback. Participants grasped a target object as quickly as possible with either a whole-hand grip or a precision grip upon hearing an auditory stimulus. The auditory cue was preceded by a visual prime stimulus. The prime depicted an object that was congruent, incongruent, or ambiguous with respect to the required grip, or it showed no object at all (no prime). Furthermore, participants performed the task in one of three vision conditions. In the "no-vision" task, participants' vision was occluded immediately after prime presentation. Consequently, participants had no vision of the target object during both the motor programming and the online control phase. In the "partial-vision" task, participants' vision was occluded as soon as they initiated their grasping movement. Hence, participants had object vision during the motor programming but not the online control phase. Finally, in the "full-vision" task, participants had vision available throughout task performance. Our results revealed the presence of two distinct priming effects. First, we found a facilitative effect of congruent compared to incongruent prime-grip combinations that influenced both reaction times and specific kinematic parameters (e.g., maximal grip aperture, peak velocity). Second, we found slower reaction times for the no prime compared to all other conditions. However, this effect was no evident in movement kinematics. Importantly, both effects were similar regardless of vision condition, indicating the prime stimuli exhibited an equal influence on memory-guided and visually guided actions. These findings further challenge the real-time view of motor programming, and indicate that visually guided and memory-guided actions revert to the same underlying representations. From a functional perspective, our findings suggest the presence of two distinct cognitive-perceptual modules related to action preparation and motor programming, and point towards possible ventro-dorsal interactions.

Interactions of Cortical Networks for Object Recognition and Object Grasping

Marc Himmelbach[1], and Mareike Gann[1]

[1]Center for Neurology, Division of Neuropsychology, Hertie-Institute for Clinical Brain Research, Eberhard Karls University, Hoppe-Seyler-Str. 3, 72076 Tuebingen, Germany

Based on the prominent two-visual-stream-hypothesis put forward by Milner and Goodale [1], object recognition on the one hand and grasping of objects on the other hand have been investigated and analyzed as separate functions in largely separated cortical systems. Obviously, our everyday experience and a number of behavioral experiments tell us that the respective cortical networks must interact and possibly overlap at some critical nodes.

In behavioral experiments with healthy participants we revealed an impact of object recognition on the implementation of movement kinematics that was denied by the latest revisions of the two-visual-stream hypothesis. We found an impact of object identity on estimations of physical object size under binocular viewing [2], [3] and revealed an influence of object recognition on grip size scaling [4].

Building on these findings we conducted an fMRI study that looked into the cortical network underlying the impact of recognition on the implementation of grasp actions. Such fMRI experiments are complicated by the fact that the mere differences in movement execution during scanning might conceal the targeted differences between different modes of input signal processing. Integrating differences in the movement kinematics into the fMRI analyses we identified a network of cortical areas that was associated with the impact of perceptual object recognition on actual movement implementation. This network combined ventral perceptual areas with dorso-ventral sensorimotor areas (Fig. 1).

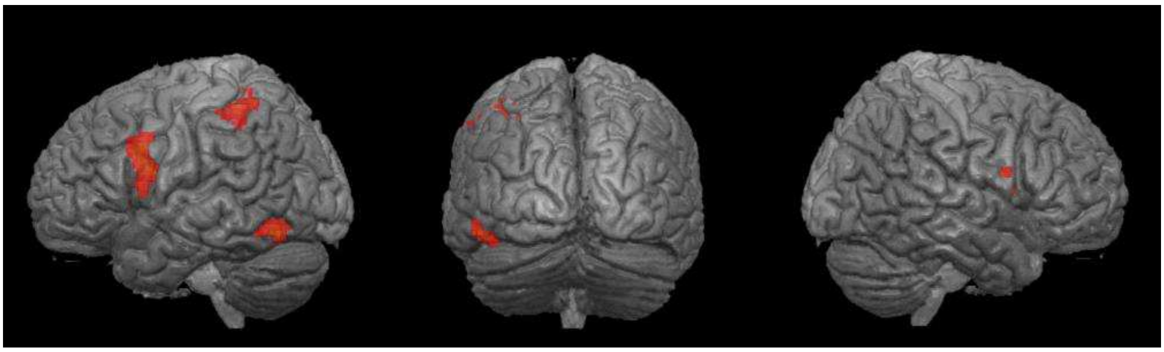


Fig. 1. Group analysis of the grasping fMRI experiment. Left IFG, left aIPS, and left lateral OTC revealed higher signal for grasping identifiable everyday objects than for featureless cuboids.

In a following fMRI study, investigating tool knowledge and the evaluation of tool efficiency, we again revealed a cortical network partially overlapping with the areas that were activated in the grasping fMRI study (Fig. 2).

Because of the fundamental differences between the two paradigms, this conjunction provides some interesting information on the functionality of these regions. In the grasping experiment the participants actually executed a movement without any overt response on object identity. In the tool evaluation experiments, no goal-directed movement was executed, but the participants gave an overt and well-considered response to a cognitive task, i.e. judging the effectivity of a particular well known or unknown tool in the context of a specific tool use task. We assume that across the two studies we identified a connected network of areas that represents a crucial site of interactions between object recognition and motor control independent of the particular output mode.

Fig. 2. Group analysis oft the tool evaluation fMRI experiment. Again, in the left hemisphere IFG, aIPS, and left lateral OTC revealedsignificant differences. However, this time between the evaluation of tool efficiency in the absence of any object-directed movements.

References

1. Milner, A. D., Goodale, M. A.: Two Visual Systems Re-Viewed. Neuropsychologia 46, 774–785 (2008)
2. Borchers, S., Christensen, A., Ziegler, L., Himmelbach, M.: Visual Action Control Does not Rely on Strangers-Effects of Pictorial Cues under Monocular and Binocular Vision. Neuropsychologia 49, 556–63 (2011)
3. Christensen, A., Borchers, S., & Himmelbach, M. (2013). Effects of pictorial cues on reaching depend on the distinctiveness of target objects. PloS One, 8(1), e54230.
4. Borchers, S., & Himmelbach, M. (2012). The recognition of everyday objects changes grasp scaling. Vision Research, 67, 8–13.

1 Symposium: Dynamics of Sketching and Sketch Understanding (DySket)

1.1 General Description:

Humans sketch and recognise (hand-drawn) sketches easily and seemingly without complex reasoning, even if the sketches were not as precise as what an artist would draw. Nevertheless, drawing sketches is an attentive process and a kind of "art" that needs to be learned, as the acquisition of sketch drawing and recognition abilities is not innate. In many cases, especially when spatial relations are of a central concern, sketches become more suitable than language and allow to more easily draw on one's well developed spatial intuitions than verbal descriptions do. Drawing on these properties, sketches are used in various ways, for example, to communicate ideas, to support design processes by externalising ideas, to understand complex relations or processes, and even to support memorisation.

Recognition or retrieval of sketches by computational tools, on the other hand, is generally difficult and requires long computations or simulation of complex mechanisms (e.g., spatial reasoning, analogy making, abstraction, etc.) that are not as intuitive as the humans' processing for sketch production or recognition.

Automating the recognition of sketches basically differs from image processing, and is a challenging task due to various reasons, particularly because of the imprecision of drawn strokes that usually constitute such sketches. Essentially, sketches are distinct from images because they transport explicit meaning, and sketch generation requires a representational apparatus that allows to design such sketches. Due to the success of touch interfaces as mainstream tools, cognitively inspired AI research faces the challenging task to develop human-computer interfaces that employ the human capacity of sketch understanding as a basis for enhanced communication with machines and modern equipment. This automation also supports human sketch usage in different use cases, such as designing early prototypes, communication, and education, to mention a few.

While clearly some pictures are sketches and some are not, it is not equally clear whether some sketches are pictures and some are not. For a broad construal of sketching and sketches, a sequence of gestures, for example, may be accepted as a sketch, while clearly not a picture. On whatever way we interpret "sketch", the worth of a sketch is determined by its quality: If a picture is worth a thousand words, does this hold also for sketches? How do humans conceptualise ideas via sketching, and how do they recognise salient parts of objects (sketched by others)? What are the main underlying cognitive mechanisms responsible for such recognition? Which parts of a sketch play more significant roles (in recognising a sketch and identifying it as a certain object) than other parts? And, more importantly, how to ultimately simulate the humans' ability to easily recognise salient concepts in sketches, and to understand what objects are sketched, in an AI model guided by the way humans operate on sketches to perform the same tasks?

The DySket symposium aims at contributing to more deeply discussing the aforementioned topics and answering questions of the latter kind on a scientific, interdisciplinary basis. The aim of the DySket symposium is to address various cognitive and computational issues ranging from applications of computational sketch understanding and generation systems, cognitive mechanisms or constraints governing sketch understanding and generation, and general principles for evaluating the quality of generated sketches and the relation of said quality to the quality of recognition. The topics of interest address various issues related to sketching and sketch understanding, as well as gestures, scene understanding and interpretation, and the generation of image schemas to depict concepts.

1.2 Speakers & Contributions:

Malumbo Chipofya *Institute for Geoinformatics, University of Münster, Germany.* (mchipofya@uni-muenster.de)

"Sketchmapia – A Framework for Recognition, Interpretation and Visualization of Sketch Maps, and Integration of Sketch Maps and Metric Maps."

This talk will give an overview of the components of Sketchmapia and one application area to which it is being applied: community based land tenure recording. The challenges imposed by the generality of free form map sketching drive us to pursue different solutions for sketch-based user interfaces for geospatial applications. We will summarize the approaches used for sketch map recognition and alignment with metric maps.

Stefan Schneider *Institute of Cognitive Science, University of Osnabrück (UOS), Germany.* (stefan.schneider@uos.de)

"Mental Object Manipulation to Generate Sketches."

While recognition of objects seemingly goes with ease, drawing an object – that is, to depict properties or perspectives of it – requires conscious effort. Two components seem to be essential: a) to be able to mentally manipulate objects, b) strategies for depicting these in 2D sketches. This talk presents results from case studies using a refined think-aloud method, where subjects acquire knowledge about geometric objects by mentally manipulating these, while solving the task to generate sketches of the constructed relations.

Oliver Kutz *Free University of Bozen-Bolzano, Italy.* (oliver.kutz@unibz.it)

"Image Schemas, Concept Invention, and Generalisation."

In cognitive science, image schemas are identified as fundamental patterns of cognition. They are schematic prelinguistic conceptualisations of events and serve as conceptual building blocks for concepts. We here propose that image schemas can also play an important role in computational concept invention, namely within the computational realisation of conceptual blending. We discuss the construction of a library of formalised image schemas, and illustrate how they can guide the search for a base space in the concept invention work flow. Their schematic nature is captured by the idea of organising image schemas into families. Formally, they are represented as heterogeneous, interlinked theories. In this context, we in particular discuss the problem of generalisation in connection with image schemas.

Kirsten Bergmann *Center of Excellence "Cognitive Interaction Technology" (CITEC), University of Bielefeld, Germany.* (kirsten.bergmann@uni-bielefeld.de)
"Social Sketching – Depicting Gestures in Multimodal Communication."
In spatial communication people spontaneously use depictive gesturing as a way to convey information. In this talk, we will present work on analyzing iconic gesture use in multimodal dialogue, their cognitive underpinnings in imagistic mental representations, and their use to establish common understanding of spatial information among communication partners. Based on empirical results, computational models are developed that allow for simulating and evaluating communicative speech-gesture behavior in artificial agents.

Zoe Falomir Llansola *Spatial Cognition Center (BSCC), University of Bremen, Germany.* (zfalomir@informatik.uni-bremen.de)
"Image Understanding Using Sketching and Qualitative Descriptors."
A computational method is presented which obtains a sketch of any digital image and then applies qualitative models (of shape, colour, topology, location, direction) to describe the features of the objects involved in that sketch. These qualitative features can be translated into narratives for human-machine interaction or into description logics for agent understanding.

Kai-Uwe Kühnberger *Institute of Cognitive Science, University of Osnabrück (UOS), Germany.* (kkuehnbe@uos.de)
"The Role of Concepts in Sketch Understanding."

1.3 Organising Committee:

Ahmed M. H. Abdel-Fattah *Faculty of Science, Ain Shams University, Cairo, Egypt.* (ahabdelfattah@sci.asu.edu.eg)

Haythem O. Ismail *Faculty of Engineering, Cairo University, Giza, Egypt.* (haythem.ismail@guc.edu.eg)

Kai-Uwe Kühnberger *Institute of Cognitive Science, University of Osnabrück (UOS), Germany.* (kkuehnbe@uos.de)

Symposium on Social Perception

Tobias Schlicht, Shaun Gallagher, and J. Suilin Lavelle

Abstract

In the debate about the nature and mechanisms of social understanding, classical mindreading approaches have been contested and supplemented by various alternatives (cf. for an overview Schlicht 2013). According to one such approach, we can directly perceive other people's mental states (Gallagher 2008). Therefore, we do not have to infer them from behavior or simulate others in order to understand what they feel and intend. Proponents of this claim argue that the classical approaches suffer from the false assumption that mental states are unobservable and hidden behind otherwise meaningless behavior. Often, this direct social perception thesis is restricted to emotions and intentions, i.e. mental states that are more closely associated with specific embodied expressions than other states like beliefs and desires. This approach has been extended to epistemic states and related to dual-process accounts of mindreading (Herschbach 2015, Apperly & Butterfill 2009) but also heavily contested on various grounds (Spaulding 2015, van Riel 2008, Jacob 2011). For example, it has been argued that a weak version of this claim that we can see mental states is compatible with versions of the mindreading approach, while a strong version is untenable (Lavelle 2012).

Social perception offers an original starting point for exploring social cognition, perception and the philosophical problem of our knowledge of other minds. Concerning social cognition, if we can visually experience others as social entities (and not as inanimate objects), then there could be a pre-conceptual appraisal of the social world that does not necessarily require conceptualization or mindreading. Appealing to perception in our appraisal of others' mental states is a first substantial step to close the gap between ourselves and others. Concerning perception, clarifying this possibility may shed light on whether we can perceive only low-level properties or also high-level properties (Siegel 2010). But is the direct social perception proposal tenable and if it is tenable, does it really provide an alternative to classical approaches? What are the assumptions and conditions of perceptual approaches?

Proponents of social perception must demonstrate (a) that mental states are perceivable and (b) clarify in which sense perception is supposed to be direct. At first sight, these two claims are independent from each other. It may be the case that a rival model, according to which perception depends on inferential mechanisms itself, as Bayesian approaches to perception suggest (Clark 2013, de Bruin & Strijbos 2015), is better suited to explain social perception. Independent from the question of whether perception is direct or not, a further issue is whether perception should be explained in enactive/embodied terms or

in representational terms and whether these approaches do in principle exclude each other (Hutto & Myin 2013, Wheeler 2008). Proponents of representational approaches to perception must clarify the nature of these representations and their intentional content, e.g. whether seeing mental states is conceptual or non-conceptual. Finally, it must be shown how the direct perception claim relates to evidence from developmental psychology about our earliest and most basic social cognitive abilities. These are some of the central questions to be addressed in this debate.

This symposium is intended as a major contribution to this debate on social perception. Many of these questions will be addressed from different philosophical perspectives and related to empirical evidence from developmental psychology and cognitive neuroscience.

Titles and Abstracts of Individual talks

Perceiving the embodied mind

Shaun Gallagher (Memphis):

Claims about the direct perception of other minds depend equally on how one characterizes perception and how one characterizes other minds. On the one hand, if one thinks of minds as internal and unobservable, then the idea of perceiving a mind doesn't make sense. On the other hand, if one conceives of the mind as embodied and situated, then in principle one can have a direct perception of the other's mind. Following this view, I'll try to say precisely what it is that we perceive when we perceive another's mind.

Which psychological states can we see?

J. Suilin Lavelle (Edinburgh):

It has been claimed by advocates of Direct Social Perception that theory-based views of mindreading must commit to the claim that minds cannot be observed. Defenders of theory-based views have responded by claiming that minds can be observed, but only if one has a 'theory of mind'. I develop a theory-based defense by exploring what kinds of psychological state can be considered observable if one is to take this approach. I will focus particularly on whether non-folk psychological states can be perceived, and, if they can, how we are to make sense of that claim.

On the nature and function(s) of social perception

Tobias Schlicht (Bochum):

Proponents of direct social perception often argue on the background of an embodied and enactive view of mental states and processes, according to which embodied expressions are constitutive elements of emotions, intentions, and possibly other mental attitudes. Enactive views typically reject classical analyses of mental states in terms of mental representations. By contrast, this talk presents arguments for a representational analysis of social perception in terms of Bayesian predictive coding, according to which perception is indirect. The focus is on the following questions: (a) What is the structure and content of the mental representations underlying social perception? (b) What is the function of social perception? – It will be shown that a representational analysis of social perception is superior to an enactive analysis. In addition, it will be shown how representational social perceptual states can play a fundamental role in the explanation of interaction and joint action, more specific social activities.

References

Apperly, Ian & Butterfill, Stephen Andrew (2009). Do humans have two systems to track beliefs and belief-like states? Psychological Review 116 (4):953.

Clark, Andy (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences 36 (3):181-204.

de Bruin, Leon & Strijbos, Derek (2015). Direct social perception, mindreading and Bayesian predictive coding. Consciousness and Cognition 36:565-570.

Gallagher, S. (2008). Direct perception in the intersubjective context. Consciousness and Cognition 17 (2):535-543.

Herschbach, Mitchell (2015). Direct social perception and dual process theories of mindreading. Consciousness and Cognition 36:483-497.

Hutto, Daniel D. & Myin, Erik (2012). Radicalizing Enactivism: Basic Minds Without Content. The MIT Press.

Jacob, P. (2011). The direct-perception model of empathy: a critique. Review of Philosophy and Psychology 2(3): 519-540.

Lavelle, Jane Suilin (2012). Theory-Theory and the Direct Perception of Mental States. Review of Philosophy and Psychology 3 (2):213-230.

Lavelle, Jane Suilin (2015). Is a modular cognitive architecture compatible with the direct perception of mental states? Consciousness and Cognition 36:508-518.

Michael, John & De Bruin, Leon (2015). How direct is social perception? Consciousness and Cognition 36:373-375.

Schlicht, T. (2013) Mittendrin statt nur dabei. Wie funktioniert soziale Kognition? In: Breyer, T. (ed.) Grenzen der Empathie. Freiburg: Alber.

Siegel, Susanna (2010). The Contents of Visual Experience. Oxford.

Spaulding, Shannon (2015). On Direct Social Perception. Consciousness and Cognition 36:472-482.

van Riel, Raphael (2008). On how we perceive the social world. Criticizing Gallagher's view on direct perception and outlining an alternative. Consciousness and Cognition 17 (2):544-552.

Wheeler, Michael (2008). Minimal representing: A response to Gallagher. International Journal of Philosophical Studies 16 (3):371 – 376.

Symposium at the meeting of the German Cognitive Science Society 2016 in Bremen

Topic of the symposium: "Mental Files in Cognitive Science: Core cognition, Concepts and Mindreading"

Speakers: Prof. Dr. Albert Newen, Philosophy, Bochum (organizer);
Prof. Josef Perner, Psychology, Salzburg; Prof. Francois Recanati, Philosophy, Institut Jean Nicod, Paris; Dr. Joulia Smortchkova, Philosophy, Bochum

General overview: Although the idea of mental files as a tool in mental representations of the human mind is not a new one in Cognitive Science, there are remarkable fruitful new usages of mental files to explain several phenomena which were waiting quite long for an adequate treatment. The symposium aims to present some of the key phenomena and include speakers who contributed to the new perspective of using mental files. The areas of application include 1. Mental Files and Concepts, 2. Mental Files in singular thought, 3. Mental Files and Theory of Mind, and 4. Mental Agency Files.

Abstracts:
Prof. Dr. Albert Newen, Philosophy, Bochum (organizer): **Introduction: Mental Files and Concepts**

One aim of the talk is to introduce the notion of mental files and a cognitive semantics as a basis for the whole symposium. Furthermore, it will be shown that concepts – which are the meaning of predicates – should best be analyzed as mental files. This is situated between empiricist accounts of concepts, which claim that concepts can be fully analyzed in terms of a network of associated perceptual information (Barsalou 1999, Prinz 2004), and rationalist accounts of concepts, which argue that concepts are radically different in format from perceptual representations (like in Fodor’s (1975) language of thought and Dretke’s (1983) theory of digital representations). I argue that a theory of concepts of mental files is able to integrate the main aspects of both the empiricist and the rationalist accounts. This can be done illustrating two main claims: 1. Concepts as mental files can be fruitfully understood as consisting of two components, (a) an *integrated network of associated information* (the empiricist part), and (b) a *handling system* that organizes this associative network (the rationalist part). 2. This account of concepts as mental files allows us to adequately describe the variety of concepts we observe in natural language (e.g. prototype concepts, definitional concepts and natural kind concepts) as well as their ontogenetic development.

Prof. Francois Recanati, Philosophy, Institut Jean Nicod, Paris; **Mental Files in a dynamic perspective**

Mental files play two main roles. First, they play the *mode of presentation* role and satisfy Frege’s Constraint : if a rational subject can believe of a given object both that it is F and that it is not F, then the subject thinks of that object under distinct modes of presentation (distinct mental files). Second, they account for coreference *de jure*. If two token singular terms a and b are associated with the same file, it is presupposed that they corefer (if they refer at all) and 'trading upon identity' becomes valid : one can move from 'a is F' and 'b is G' directly to 'there is an x which is F and G', without needing to invoke an identity premiss (Campbell 1987). It has been argued that there is a tension between the two roles, and that a single entity cannot play both. I will show that the tension can be alleviated by distinguishing mental files as continuants (dynamic files) and time-slices thereof (static files).

Josef Perner, Centre for Cognitive Neuroscience and Department of Psychology, University of Salzburg: **Mental Files Theory of Mind**

I provide a cognitive analysis of how we represent belief using mental files. For each relevant object a regular file represents one's own view and a coreferential vicarious file captures another person's beliefs about the object. This analysis enables to predict and explain several developmental phenomena. Around 4 years children pass the false belief test as they start to understand identity statements and alternative labelling of objects. However, it takes another 2 years for children to master the *intensionality* or *aspectuality* of knowledge and belief, i.e., that it matters under which description a believer is acquainted with an object, e.g., that she knows the die/eraser as a die but not as an eraser. This understanding is achieved as children become able to comprehend second-order embedded states, e.g., she *thinks* she *knows*. The theoretical analysis of these achievements is supported by empirical results.

Joulia Smortchkova, Philosophy, Bochum: **Core agency cognition: from object-files to agent-files**

Infants younger than one-year-old show an innate understanding of how inanimate objects behave. For example, they do not expect objects to change shape, to interpenetrate or to disappear forever when traveling behind a barrier (Spelke, 1990). Developmental psychologists suggest that the "core object system" underpins these abilities (Carey, 2009). This object tracking system persists in adulthood: adults track visual objects via mid-level visual representations called "object-files" (Scholl, 2001).

Building on a recent proposal by Murez and Smortchkova (2014), I extend the notion of "object-files" to "agent-files", i.e. visual representations within the "core agency system" dedicated to tracking animate entities. While similar to object-files in many respects, the cognitive roles and triggering and maintenance conditions of agent-files are nevertheless significantly different from those of object-files.

In my presentation, I argue that agent-files ground a distinction between inanimate objects and animate entities within perception. Agent-files thus play a crucial role in infants' early appraisal of the social world.

References for all talks (speakers highlighted in bold):

Barsalou, L.W. (2012): The human conceptual system. In: *Cambridge Handbook of Psycholinguistics*, hg. v. M. Spivey, K. McRae & M. Joanisse. Cambridge, MA; S. 239-258.

Carey, Susan (2009). *The Origin of Concepts*. Oxford University Press.

Murez, Michael & **Smortchkova, Joulia** (2014). Singular Thought: Object-Files, Person-Files, and the Sortal PERSON. *Topics in Cognitive Science* 6 (4):632-646.

Newen, A. (2010/11). Phenomenal Concepts and Mental Files: Phenomenal Concepts are theory-based. In: *philosophia naturalis* 47/48, S. 155-183.

Newen, A, Marchi, F. (2016). Concepts and their organizational structure: Concepts are templates based on mental files. In: D. Hommen, C. Kann, T. Osswald (Hrsg.): *Concepts and Categorization. Systematic and Historical Perspective*, mentis: Münster, S. 197-226.

Perner, J., Huemer, M., & Leahy, B. (accepted 11 August 2015). Mental Files and Belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*.

Perner, J. & Leahy, B. (published online 22 March 2015). Mental files in development: Dual naming, false belief, identity, and intensionality. *Review of Philosophy and Psychology*, 6, pp-pp. DOI: 10.1007/s13164-015-0235-6

Perner, J. & Roessler, J. (2012). From Infants' to Children's Appreciation of Belief. *Trends in Cognitive Sciences*, 16 (10), 519–525.

Perry, J. (1990): Self-Notions. In: *Logos* 11, S. 17-31.

Recanati, F. (2012a): *Mental Files*, Oxford: *Oxford University Press*.

Recanati, F. (2012b). Empty Singular Terms in the Mental-File Framework. In. M. Garcia-Carpintero and G. Marti (Eds): Thinking and Speaking about Nothing. Oxford: *Oxford University Press*.

Spelke, Elizabeth S. (1990). Principles of object perception. *Cognitive Science* 14 (1):29--56.

Insight and evolution

Anna Fedor [1,3], István Zachar [2,3], András Szilágyi[1,3], Michael Öllinger [3], Harold P. de Vladar[3], Eörs Szathmáry[1,2,3]

[1] MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Budapest, Hungary

[2] Department of Plant Systematics, Ecology and Theoretical Biology, Institute of Biology, Eötvös University, Budapest, Hungary

[3] Parmenides Center for the Conceptual Foundations of Science, Munich/Pullach, Germany

Abstract

In this talk I will show our most recent results which show that evolutionary processes can solve insight problems.

The proposed model, called Darwinian Neurodynamics (Fernando, Goldstein, and Szathmáry 2010; Fernando and Szathmáry 2009, 2010), is implemented as an artificial neural network, where attractor networks store and reproduce candidate hypotheses. Neural Darwinism is a hypotheses that has been established decades ago (Changeux, Courrège, and Danchin 1973; Edelman 1987). It explains how selective pruning shapes the brain during development. However, it does not explain, how new ideas are generated and it is described as only one round of selection. Darwinian Neurodynamics is more powerful in the sense that it allows selection in several rounds and it also explains the generation of new variants. We suggest that evolutionary processes play a role in human cognition and our attractor network-based implementation is a plausible model for certain aspects of problem solving.

We chose a well-known insight problem, the four-tree problem to demonstrate how Darwinian Neurodynamics solves problems. The instructions for the four-tree problem are the following: *A landscape gardener is given instructions to plant four special trees so that each one is exactly the same distance from each of the others. How is he able to do it?* (de Bono 1967). The solution is that he plants the trees on the apices of a regular tetrahedron, which is easiest to do if one of the trees is on top of a hill, and the other three trees are at ground level in a shape of a triangle. We simplified the task, so that it is reduced to only planting the fourth tree; the other three trees are already planted in a shape of a triangle. The task of the model is to find the position of the fourth tree in a three-dimensional coordinate system: each candidate solution unambiguously defines the proposed coordinates of the fourth tree. Candidate solutions are represented by the activation patterns of the attractor networks in the system.

Our results show that the model can solve the task, moreover, it behaves similarly to human problem solvers. When it is initialized in a way that is conceptually similar to how human problem solvers receive the task, it starts to search for the position of the fourth tree in two-dimensional space and after switching from selection to evolution mode, it starts to search in three-dimensions – a behavior that mimics representational change. Its solution rate can be increased by pretraining it and by priming it with three-dimensional patterns, just like in experiments with human participants (Kershaw, Flynn, and Gordon 2013).

Our work shows that evolutionary processes can solve the four-tree problem, and that our model behaves comparably to human participants. Future work should include experimentally testing the predictions of the model and to implement a more realistic coding of the task.

References

de Bono, Edward. 1967. *The Use of Lateral Thinking*. London: Jonathan Cape.

Changeux, J P, P Courrège, and a Danchin. 1973. “A Theory of the Epigenesis of Neuronal Networks by Selective Stabilization of Synapses.” *Proceedings of the National Academy of Sciences of the United States of America* 70(10): 2974–78.

Edelman, G M. 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Fernando, Chrisantha, Richard Goldstein, and Eörs Szathmáry. 2010. “The Neuronal Replicator Hypothesis.” *Neural Computation* 22(11): 2809–57.

Fernando, Chrisantha, and E Szathmáry. 2010. “Chemical, Neuronal and Linguistic Replicators.” In *Evolution—the Extended Synthesis*, MIT Press, 209–50.

Fernando, Chrisantha, and Eörs Szathmáry. 2009. “Natural Selection in the Brain.” In *Towards a Theory of Thinking: Building Blocks for a Conceptual Framework*, eds. Britt Glatzeder, Vinod Goel, and Albrecht Müller. Berlin, Heidelberg: Springer Berlin Heidelberg, 291–322.

Kershaw, TC, CK Flynn, and LT Gordon. 2013. “Multiple Paths to Transfer and Constraint Relaxation in Insight Problem Solving.” *Thinking & Reasoning* 19(1): 96–136.

On the role of physical space for commonsense problem solving

Christian Freksa

University of Bremen

Classical artificial intelligence (AI) employs logical reasoning as the foundation of cognitive processing. An early research field within AI has been *commonsense reasoning* whose approaches have been built on this foundation. If the logic-based models of the commonsense reasoning community are intended as role models of cognitive processing, their creators imply that commonsense reasoning is based on logical reasoning abilities.

In a recent article, Davis and Marcus (2015) point out that progress in commonsense reasoning has been much slower than in other areas of AI. I would like to suggest that one reason for this might be that logic is an inappropriate substrate for commonsense capabilities, as logical reasoning abilities follow rather than precede commonsense abilities.

In animal (including human) development spatial abilities precede logical reasoning abilities; therefore I propose that the laws of physical space might serve as a better foundation of commonsense capabilities than the laws of formal logic.

Suppose we want to build a robot that can solve spatial puzzles, such as shape sorting, snake cube, the 15-puzzle, or the shortest route problem. There are a variety of qualifications and skills of cognitive agents that may be useful for solving spatial puzzles, including manual skills, imaginativeness, attentiveness, re-representation skills, analytical skills, reasoning skills, perceptual sharpness, and curiosity. There also are a variety of circumstances that may be helpful for solving spatial puzzles, including the spatial starting configuration, perceptibility of the spatial configuration, manipulability of the configuration, similarity of the puzzle to known problems.

In my contribution I want to discuss some of these qualifications, skills, and circumstances as well as their interrelationships. Based on these considerations I will address the question of how we could conceive and develop a robot from scratch that is capable of solving spatial puzzles.

Spatial puzzles have been selected as well-structured examples of basic everyday challenges cognitive agents may be confronted with. They require spatial cognitive abilities – a combination of physical and mental abilities. Spatial cognitive abilities may be forerunners or prerequisites of more abstract cognitive abilities. In this sense, constructing spatial puzzle solvers is viewed as developing a theory of cognitive abilities on the basis of the ecology of interactions in space.

Experience, Understanding and Creativity

Bipin Indurkhya
Jagiellonian University, Cracow, Poland
bipin.indurkhya@uj.edu.pl

1 Experience and Creativity

We outline a framework of cognition and perception that makes a distinction between experience, understanding and explanation, as shown in Fig. 1. Experience refers to immediate perceptual awareness. Understanding occurs when the experience is incorporated into some internal model, by relating to past experiences etc. Experiences and understanding are largely subjective. We get to the explanation level when the experience and the understanding are formulated in a language or some other communication medium so that it can be shared. Thus, experience is understood, and understanding is communicated through explanation.

The key point of the model is that the flow of information between these three levels proceeds in both directions. So though experience leads to understanding, prior understanding effects experience. Similarly, though understanding leads to explanation, prior explanation effects understanding. Often, the influence of experience on understanding, and the influence of understanding on explanation, is referred to as bottom-up processing, and the influence of explanation on understanding, and the influence of understanding on experience, is referred to as top-down processing.

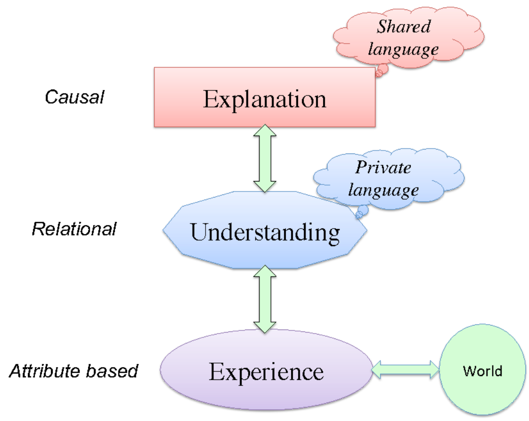


Fig. 1: A model of perception and cognition

This is known as the interaction model of cognition in that the conceptual understanding of the world (or the environment) is seen to arise as a result of interaction between concepts and percepts. The origin of this view can be traced to Kant, who argued against purely empirical accounts (purely bottom-up processing), such as John Locke, and also against purely conceptual accounts (purely top-down processing), such as Plato, to conclude that concepts without percepts are empty, and percepts without concepts are blind. Later on, Cassirer argued that a multiplicity of symbolic systems (or worlds) is possible, and that they might not be reducible to one another. (See also Goodman 1978.) Gestalt psychologists showed that mind imposes its own organization on the perceptual stimuli. Piaget provided further evidence that cognition is a constructive process in which a cognitive agent structures the world based on its actual and potential actions. (See also Indurkhya 1992; O'Regan 2001).

More and more evidence has emerged to show that one's beliefs and expectations have a remarkable influence on one's perception. For example, Lang *et al*. (1975) showed that the aggressive behavior of participants after drinking was affected by what they thought they were drinking rather that what they actually drank. Brochet (2001) found similar effect with wine tasting: the participants' perception of taste depended on the bottle from which the wine was poured, even though it was the same wine. All this points to the fact that our expectations, past experience, and understanding influences our perception.

2. Conceptualization and loss of information

As we go from experience to understanding to explanation, there is invariably a loss of information. This is because an experience can be understood in many ways, and an understanding can be explained in different ways. So once a particular understanding is chosen, and a particular explanation is chosen, the alternative ways for understanding and explanation are lost. This loss of information can be understood in the following way. When an object is described as a 'chair', many specific properties of that chair, like its color, style, kind of material etc. are lost. Of course, we could make our conceptualization of the object more specific — it is a red chair, made of teak, with a high back, and so on — but no matter how detailed the conceptual representation is made, there is always some aspects of the object that are excluded, and it is these excluded aspects that constitute the *information lost in the conceptualization*. (This precisely is the theme of a short story *Del Rigor en la Ciencia (On Exactitude in Science)*, by Jorge Luis Borges and Adolfo Casares.)

3. Crux of Creativity: Re-conceptualization

In the model of cognition we have sketched above, we can say that most of the time when we have an experience, we habitually understand in a certain way; and given an understanding, we habitually explain it in a certain way. This comes from the top-down influences that reflect cultural and habitual conditioning. This is graphically shown in Fig. 2.

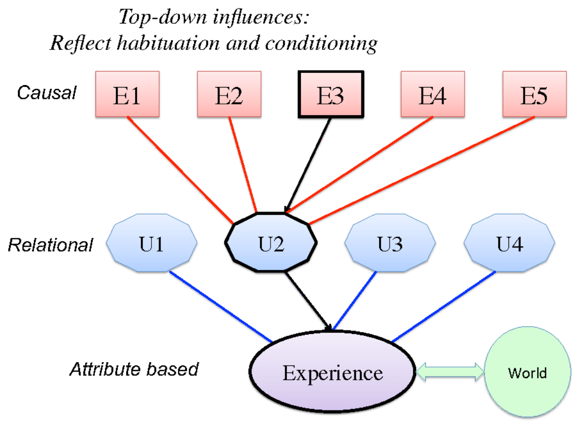


Fig. 2: Top-down influences: Reflect habituation and conditioning

The information that is lost in this habitual conceptualization is something that was perhaps not relevant to the cognitive agent (or to the society of the cognitive agents) in their past interactions. However, this information may become crucial to solve a new problem, and then the habitual understanding and explanations become hopelessly inadequate. Creativity, then, lies in alternate ways to understand an experience; and alternate ways to explain an understanding. In order to succeed, it must undo the effects of top-down conditioning, move closer to perception, and focus on the experience (Fig. 3). Needless to say, this is a cognitively difficult task, for we are strongly conditioned by our culture and habits, and it is not so easy to break away from it. Approaches such as meditation that focus on sensation, art lessons that focus on seeing or listening, metaphors that focus on inter-domain or cross-modal connections, are all designed to help overcome this bias and conditioning.

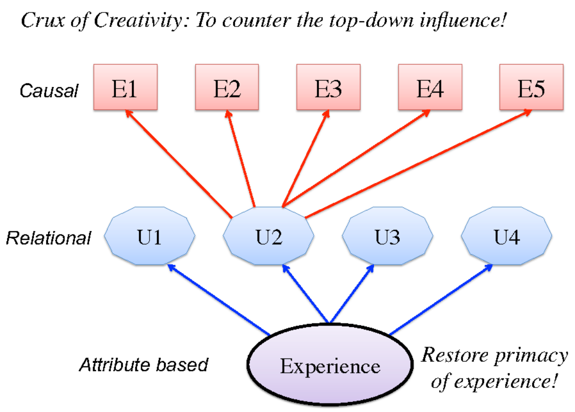


Fig. 3: Crux of Creativity: To counter the top-down influence!

This process of re-conceptualization allows one to recover some of the information that was lost in the habitual understanding and explanation. This can be useful for aesthetic pleasure, in art and poetry; for making new artifacts, in science and technology; or for achieving some desired situation, in problem solving. This, in the model we are proposing here, is the crux of creativity. (See also Indurkhya 1999; 2006; 2010.)

References

1. Brochet, F.: *Tasting: Study of chemical representations of objects in the field of consciousness*. PhD thesis. Faculty of Oenology of Bordeaux. Laboratory of General Oenology. (2001).
2. Goodman, N.: *Ways of worldmaking*. New York: Hackett (1978).
3. Indurkhya, B.: *Metaphor and Cognition*. Kluwer Academic Publishers, Dordrecht, The Netherlands (1992).
4. Indurkhya, B.: Emergent representations, interaction theory, and the cognitive force of metaphor. *New Ideas in Psychology 24* (2), 133–162 (2006).
5. Indurkhya, B.: On the role of metaphor in creative cognition. *Proc. International Conference on Computational Creativity: ICCC-X,* Lisbon, Portugal, 51–59 (2010).
6. Lang, A.R., Goeckner, D.J., Adesso, V.J., Marlatt, G. A. : Effect of alcohol on aggression in male social drinkers. *Journal of Abnormal Psychology 84*(5), 508–518 (1975).
7. O'Regan, J.K., Noë, A.: A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–1011 (2001).

Towards Re-representation in Cognitive Systems

Ana-Maria Olteţeanu[1]

Cognitive Systems, Bremen Spatial Cognition Center,
Universität Bremen

1 The role of re-representation in creative problem solving

The ability to re-represent problems has been long acknowledged as a component of creative problem solving, in the case of insight problems [1, 4]. In this case, the ability for re-representation generally means an ability to internally or externally represent the problem in a way which will afford the solution with more ease. Saying that a problem has been re-represented might mean, amongst others:
(i) that elements of the problem which were not salient have become so, (ii) that the required solution or goal has itself been represented in different ways, or
(iii) that different structure is being used to organize the problem [5] - a structure which has already proven to have some solution in the past.

According to this, a definition of problem solving accounting for creativity – thus a definition of *creative problem solving* [6] – would have to amend the classical definition of problem solving in the following ways:

- Before exploring the problem space from some initial state, a problem representation is constructed;
- This problem representation is in some measure arbitrary and not the only representation which is possible (not unique) - reflecting perhaps the closest associative knowledge that the problem solver brings to bear;
- This representation yields an initial state and a problem space, with attached operators and paths (the operators known or strongly associated by the problem-solver to those representations); problem solving may proceed henceforth in the classical way;
- If the problem solving process is not successful, it can restart with more or less ease (depending on creative ability) from the step of representing the problem; this will yield different associated problem spaces;
- The representation of the problem can be changed in a variety of ways, including but not limited to (i) bringing new features and objects into the problem; (ii) using different known representations to parse the problem and (iii) changing the currently held representations.

In this talk we will aim for a better definition of re-representation processes, as encountered in natural cognitive systems, so that such an understanding can be applied to implementations of artificial cognitive systems capable of re-representation. The talk will focus in turn on (a) exemplifying re-representation which may be spatial in nature in the context of insight problems, (b) showcasing possible multiple levels of re-representation in the context of different tasks, and allthroughout these examples (c) clarifying some of the parameters which could support the implementation of such processes in artificial cognitive systems.

2 Re-representation with spatial underpinnings in insight problems

Different representations of various insight problems can be shown to emphasize or de-emphasize various problem elements, making problems easier or harder to solve. In this part of the talk we will describe a few such representations, their consequences, and the links of re-representation to restructuring.

An example of a problem which can be shown in such re-representational cases is the chain problem [2]. The chain problem, shown in Fig. 1 is defined as follows: *A girl has four pieces of a chain (see Figure 1). Each piece is made up of three links. She wants to join the pieces into a single closed loop of chain (like a necklace). To open a link costs 2 cents and to close a link costs 3 cents. She only has 15 cents. How does she do it?*

Fig. 1. The chain problem

A possible re-representation of this problem, illustrated in Figure 2, is the following: *You are given four sets of zeroes. You must make one complete line of zeroes, with no gap, and also cover the initial x. You are allowed to move the zeroes on the board. You can move: (i) either one zero at the time, wherever you like, or (ii) an entire set of three zeroes. Each such move costs you 5 points, you only have 15 points you can use.*

Fig. 2. The chain problem re-represented

The second representation of what is actually the same problem is generally perceived as a much easier problem. Such types of re-representation, and the links between re-representation and restructuring are discussed.

3 Different levels of re-representation

Re-representation might be an element encountered not only at the level of insight problems, but in all sorts of different types of tasks too. Three other types of tasks are proposed as accounting for such levels, besides insight problems: ambiguous figures reversal, interpreting meaning of ambiguous pattern stimuli in a Wallach Kogan test [7] and seeing objects as possible other objects in the Alternative Uses test [3]. The fourth type of task is constituted by insight problems of various kinds: e.g. abstract problems which allow restructuring, like the ones above, and practical object use problems.

Alltogether, these constitute various levels of feature, object and problem representation, as follows:

a) representation and re-representation of features as various feature subsets of various objects in figure reversal and ambiguous pattern stimuli;
b) representation and re-representation of objects as different objects in the process of coming up with creative object uses;
c) representation and re-representation of objects and problems as different objects and problem structures in the context of insight problems.

References

1. Batchelder, W.H., Alexander, G.E.: Insight problem solving: A critical examination of the possibility of formal theory. The Journal of Problem Solving 5(1), 56–100 (Fall 2012)
2. Dow, G.T., Mayer, R.E.: Teaching students to solve insight problems: Evidence for domain specificity in creativity training. Creativity Research Journal 16(4), 389–398 (2004)
3. Guilford, J.P.: The structure of intellect. Psychological bulletin 53(4), 267 (1956)
4. Ohlsson, S.: Restructuring revisited: II. An information processing theory of restructuring and insight. Scandinavian Journal of Psychology 25, 117–129 (1984)
5. Olteţeanu, A.M.: Two general classes in creative problem-solving? an account based on the cognitive processes involved in the problem structure - representation structure relationship. In: Besold, T., Kühnberger, K.U., Schorlemmer, M., Smaill, A. (eds.) Proceedings of the Workshop Computational Creativity, Concept Invention, and General Intelligence. Publications of the Institute of Cognitive Science, vol. 01-2014. Institute of Cognitive Science, Osnabrück (2014)
6. Olteţeanu, A.M.: Seeing as and re-representation: Their relation to insight, creative problem-solving and types of creativity. In: Besold, T., Kühnberger, K.U., Schorlemmer, M., Smaill, A. (eds.) Proceedings of the Workshop on Computational Creativity, Concept Invention, and General Intelligence. Publications of the Institute of Cognitive Science, vol. 02-2015. Institute of Cognitive Science, Osnabrück (2015)
7. Wallach, M.A., Kogan, N.: Modes of thinking in young children: A study of the creativity-intelligence distinction. Holt, Rinehart & Winston (1965)

Challenges and Directions for Making Cognitive Systems Creative

Kai-Uwe Kühnberger

Institute of Cognitive Science
University of Osnabrück
Albrechtstr. 28, Germany
kkuehnbe@uos.de

Abstract. In this paper, we will roughly sketch the idea of cognitive mechanisms like analogy-making and conceptual blending as means for modeling and explaining creative abilities of humans. Then, we will discuss IBM's Watson/Bluemix services, which are intended to go beyond classical computing paradigms in order to approach the level of "cognitive computing" in its human-inspired facets. We will discuss challenges that arise if systems like Watson/Bluemix are used for implementing creative abilities. Finally, some speculations about possible directions for addressing these challenges in cognitive systems by referring to cognitive mechanisms and their corresponding computational realizations will be mentioned.

1 Analogy-Making and Conceptual Blending as a Source for Creativity

Modeling creative abilities with computing devices is considered to be a hard problem. Despite difficult questions like what creativity is, whether creativity should be assigned to a cognitive process, to the product of such a process, or to both, or how creativity can be assessed and measured, there is a strong interest in the last years to develop computational approaches for creativity. Some of these theories focus on the modeling of cognitive mechanisms like analogy-making and conceptual blending. By applying analogy-making and conceptual blending to problem solving tasks, it turns out that both mechanisms are useful to model certain creative abilities of humans. Examples of such applications can be found in fields like problem solving [2], mathematics [5], music [1], or physics [8].

Analogy-making can be understood as the detection of structural commonalities of two domains [7]. Domains can be a variety of things, e.g. conceptual spaces, micro-theories, or representations of commonsense knowledge. Some researchers claim that the ability to establish analogical relations is a core of human cognition [4]. On the other hand, conceptual blending in the sense adopted here takes two input spaces and attempts to compute a generic space and a blend space, i.e. the latter being a new and independent conceptual space containing a mixture of conceptual information from both input domains [2].

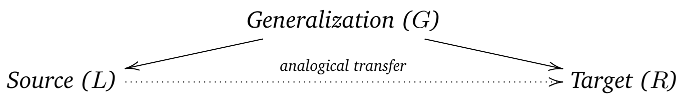


Fig. 1. HDTP's overall approach to creating analogies (cf. [8]).

Fig. 1 depicts the general idea of the analogy engine Heuristic-Driven Theory Projection (HDTP) [8]. Given a source S and a target T, the HDTP engine computes an analogical relation between S and T together with a generalization G that covers the structural commonalities between S and T. The conceptual idea of conceptual blending in Joseph Goguen's work is slightly similar to the analogy-making process (compare Fig. 2): in a blending process two input spaces S and T are generalized as well as merged in a blend space B. We consider conceptual blending therefore as a twofold process, which first detects structural similarities based on an analogical relation between S and T and then merges the two input spaces based on appropriate heuristics.

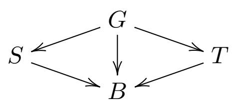


Fig. 2. Goguen's version of concept blending (cf. [3]).

2 Remarks on IBM's Watson/Bluemix Services

An industry system that is claimed to go beyond classical computing paradigms in order to approach "cognitive computing" or the "cognitive era of computing" in its human-inspired facets is IBM's Watson/Bluemix platform. Besides classical industry applications that emphasize the integration of Big Data with highly structured knowledge in tasks such as predictive maintenance, production processes, and recommendation systems for health applications, IBM applied the Watson/Bluemix services also in the domain of creativity research. An example is "IBM Chef Watson" (cf. https://www.ibmchefwatson.com), where innovative culinary recipes are generated by the system. A part of the scientific background of the approach can be found in [6].

The strength of IBM's system is based on the combination and integration of different types of knowledge resources and the possibility to combine a large number of different algorithms including deep learning theories. For example, the Watson services that are currently offered range from classical NLP applications (like the conversation of speech to text and vice versa), to such services like AlchemyAPI (service to extract semantic meta-data from images), concept insights (used for content exploration and recommendation of texts), or personality insights (classification of people's personality characteristics based on textual descriptions) just to mention a few of them. For such services different algorithms are used and the strength of the overall architecture can be seen in the possibility to combine these services.

An exemplification of such a combination of different knowledge resources and algorithms is the Flu Prediction study project that took place recently in Osnabrück (http://www.flu-prediction.com/about). The task is to improve the prediction of influenza spreading across the US. The approach combines data science methods to allow the identification of complex causal relations, the efficient use of Watson/Bluemix services, the usage of social media data (Twitter) and conventional data from the Center for Disease Control (CDC), and Watson as a question-answering system to allow the user to receive background information about the flu.

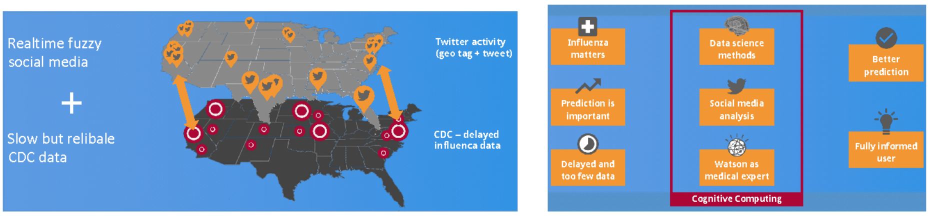


Fig. 3. The left image depicts a graphical representation of the combination of reliable but delayed CDC data about influence and current social media data (based on real-time Twitter activity). The image on the right side shows the overall idea to combine services and algorithms to get better predictions about the future spreading of the flu. Both images are taken from an online lecture given by Gordon Pipa and available on Youtube: https://www.youtube.com/watch?v=I2FcCpXwxVw

Although the strength of the described combination of services for cognitive computing applications is for certain domains quite remarkable, a natural question is whether this suffices to develop strong models of computational creativity as well. The application "IBM Chef Watson" mentioned above seems to allow such an extension to creativity tasks, nevertheless, further examples are rare and the general problem remains whether systems based on Big Data can be used for highly abstract domains like the invention of new concepts in mathematics or physics. Furthermore, to assess the degree to which approaches similar to IBM's Watson/Bluemix system can be called cognitively inspired is less clear given that classical cognitive mechanisms like analogy-making, conceptual blending, similarity, or heuristic assessments play only, if at all, a very limited role. Finally, the cross-domain aspect of cognitive mechanisms, namely the ability to transfer certain patterns from one domain to a totally unrelated domain, in order to solve a certain problem as well as the multi-modal knowledge integration ability of humans is often missing in classical cognitive systems.

3 Possible Directions for Addressing These Challenges

In order to sketch ideas how to extend cognitive computing systems to improve their creative capacities, we propose the following extensions.

– Models of computational creativity in highly abstract domains like mathematics require more than visual or textual information. The natural choice for a representation of mathematical knowledge are axiomatic systems formulated in a first-order or better monadic second-order language. A natural choice to expand cognitive systems such as IBM's Watson/Bluemix system would be to add services that can deal with such axiomatic theories. In [5], it is shown how reasoning on such representations can be used for modeling creative processes.
– Humans use multi-modal representations for all sorts of cognitive tasks. For example, verbal communication of humans does neither only work on the syntactic level of natural language, nor only on the semantic level. Verbal communication is rather a complex interplay between linguistic representations (syntactic, semantic, pragmatic etc.) and non-verbal means of communication like gestures, expression of emotions, motor actions etc. Similarly, in solving creativity tasks, humans often change representations between abstract and concrete representations, switch between and integrate multi-modal information, and project knowledge between these representation types. A creative cognitive system should also be able to use such multi-modal representation. In the Watson/Bluemix system, a first step into this direction is the multi-modal integration of image and text.
– In order to make cognitive computing *cognitive* in a strong sense, cognitive mechanisms such as analogy-making and conceptual blending should be added to such systems. It was argued in Section 1 that a large variety of creativity aspects in various domains can be computationally modeled by implementations of such cognitive mechanisms. Expanding cognitive systems into such a direction has the potential to increase the capabilities of such systems significantly.

References

1. M. Eppe, R. Confalonieri, E. Maclean, M. Kaliakatsos-Papakostas, E. Cambouropoulos, M. Schorlemmer, M. Codescu & K.-U. Kühnberger (2015): Computational Invention of Cadences and Chord Progressions by Conceptual Chord-Blending, IJCAI 2015.
2. G. Fauconnier & M. Turner (2002): The Way We Think. Conceptual Blending and the Mind's Hidden Complexities, Basic Books, New York 2002.
3. J. Goguen (2006): Mathematical models of cognitive space and time. In: D. Andler, Y. Ogawa, M. & S. Watanabe (eds.): Reasoning and Cognition: Proc. of the Interdisciplinary Conference on Reasoning and Cognition, Keio University Press.
4. D. Hofstadter (2001): Epilogue: Analogy as the core of cognition. In: D. Gentner, K. Holyoak, B. Kokinov (eds.): The Analogical Mind: Perspectives from Cognitive Science, MIT Press, pp. 499–538.
5. M. Martinez, A. Abdel-Fattah, U. Krumnack, D. Gomez-Ramirez, A. Smaill, T. Besold, A. Pease, M. Schmidt, M. Guhe & K.-U. Kühnberger (2016): Theory blending: extended algorithmic aspects and examples. Annals of Mathematics and Artificial Intelligence, DOI: 10.1007/s10472-016-9505-y.
6. F. Pinel, L. Varshney & D. Bhattacharjya (2015): A Culinary Computational Creativity System. In: T. Besold, M. Schorlemmer & A. Smaill (eds.): Computational Creativity Research: Towards Creative Machines, Amsterdam, Paris, Beijing, Atlantis Press, pp. 327–346.
7. M. Schmidt, H. Gust, K.-U. Kühnberger & U. Krumnack (2011): Refinements of Restricted Higher-Order Anti-Unification for Heuristic-Driven Theory Projection. In: KI 2011: Advances in Artificial Intelligence, LNCS vol. 7006, pp. 289–300.
8. A. Schwering, U. Krumnack, K.-U. Kühnberger & H. Gust (2009): Syntactic Principles of Heuristic-Driven Theory Projection. Cognitive Systems Research 10(3):251–269.

Qualitative reasoning models to help solving spatial ability tests

Zoe Falomir

Spatial Cognition Centre, Universität Bremen, Germany
zfalomir@informatik.uni-bremen.de

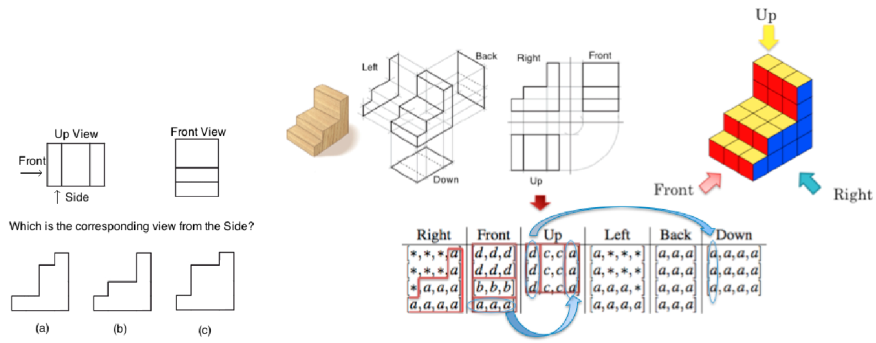
1 Extended Abstract

Spatial cognition studies showed that there is a strong link between success in Science, Technology, Engineering and Math (STEM) disciplines and spatial abilities [12, 18]. These abilities are basic for job skills like visualizing the result of a surgery, designing bridges, aircrafts; interpreting charts, maps, engineering drawings, etc. Moreover, 3D spatial skills can be developed through practice since research [15] showed that students who attended a course at university to improve their ability to visualize in 3D, improved their success and retention significantly, particularly female students. Thus, it is important to train spatial abilities from the early stages, which can have a beneficial impact on gender equality. The actualities of training spatial reasoning also in contemporary school mathematics are studied in US and Canada [14].

Artificial agents with strong intelligence must have reasoning mechanisms to solve spatial problems cognitively. Qualitative representations are thought to be close to the cognitive domain, as shown in cognitive models of sketch recognition [9], spatial problem solving tasks (i.e. visual oddity tasks) [10] and in mental rotation tasks [11]. Novel models which combine qualitative models, cognitive spatial thinking and common sense are a challenge which envision further advances in Artificial Intelligence and its applications. The research presented here is aimed towards solving the challenge of defining models which can:

– help people to understand how to solve perceptual ability tests (i.e. paper folding and perspective tests), so that they can improve their spatial cognition skills and therefore enhance their success in STEM; and also,
– be used by artificial intelligent agents to solve spatial problems, so that they can learn spatial transformations happening when folding a paper or when seeing an object from different perspectives and then developing a new framework in spatial reasoning.

Qualitative models that try to solve spatial cognition problems have appeared in the literature [5, 4]. A qualitative model for describing 3D objects (Q3D) using depth and different perspectives [5] was inspired by designs of pieces which abstract the main features of objects from all their properties in the real world and describe them using 3 canonical views (top, lateral and front) since, in experimental psychology, there is support for the general idea that human object recognition involves view-dependent representations, that is, people prefer to imagine, view, or photograph objects from certain "canonical" views [13]. The Q3D was also motivated by the fact that: (i) the German Academic Foundation uses consistent view/projection of a 3D object corresponding to a technological drawing to measure intelligence in students[1]; and (ii) the Dental Admission Testing Program[2] by the American Dental Association (ADA) includes a Perceptual Ability Test (PAT) which part-2 is a 3D object perspective reasoning test. Fig. 1 shows an example of a perspective question and the relations between the perspectives that the Q3D model must maintain to be consistent with real space. The Q3D approach



(a) Question (b) Q3D model

Fig. 1. Q3D model: (a) example of a 3D perspective question and (b) the relations between the depth and the sketches (red lines) and the depth relations between perspectives (arrows).

allows reasoning through logics defined to test the consistency of descriptions. Q3D infers some features of the unknown perspectives of the object (i.e. back, down and the other lateral) by defining logics based on the continuity of holes and the relative depth presented by opposite and neighbouring perspectives.

A Qualitative Descriptor for Paper Folding (QD-PF) was defined [4] by establishing a correspondence between the possible folding actions and the areas in the paper where a hole can be punched. This QD-PF is motivated by the fact that paper folding tasks have been extensively used in psychological cognitive tests to measure people spatial abilities as a form of intelligence, for example, by using the *Manual for Kit of Factor-Referenced Cognitive Tests* [3], which included a factor kit Visualization (Vz) part in which Paper Folding Test (suggested by Thurstone's Punched Holes [2]) is part 2 (Vz-2). Moreover, the PAT by the ADA also includes paper folding and punching in part 4. In the QD-PF [4], reasoning tables were defined for inferring the location equivalences encountered after a paper is folded. Using these tables, logic solutions were obtained to find out where the hole punched is replicated after one-to-three foldings of a paper. Fig.

[1] *Test der Studienstiftung: Gehirnjogging für Hochbegabte*, see Spiegel Online: http://www.spiegel.de/quiztool/quiztool-49771.html

[2] *Dental Admission Testing Program example*: http://www.ada.org/

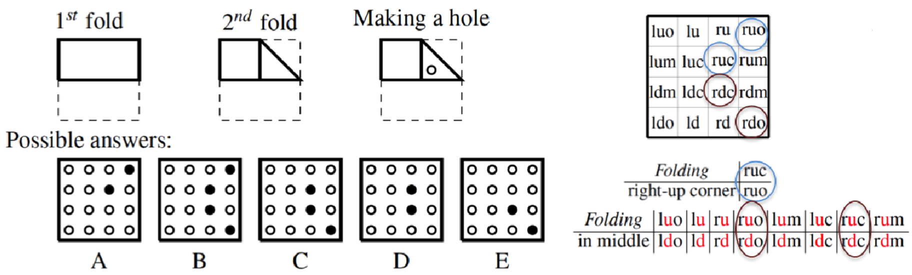


Fig. 2. QDPF model: Example question in the *Paper Folding Test - VZ-2* [3] and the inference tables that provide the logical answer to the problem. Note that the correct answer is B.

2 shows an example of a paper folding question and how the reasoning tables in QD-PF can obtain the answer.

In order to test the models Q3D and QD-PF, we selected Prolog programming language [17], which is based on Horn clause logic [8]. SwiProlog[3] was the testing platform. The testing dataset for Q3D was extracted the PAT (part-2) by ADA. The testing dataset for QD-PF was the *Paper Folding Test - VZ-2* from the *Manual for Kit of Factor-Referenced Cognitive Tests*[3].

Research has also showed that video game training enhances cognitive control [16] specially when aging [1] and that realistic 3D views enhanced users' performance on spatial visualization tests [19]. Thus, a computer game showing realistic 3D views of objects was created for testing the Q3D model. The Q3D-Game [7, 6] was developed using Unity 5 game engine and built in an Android tablet for providing interactivity with the user. The Q3D-Game prototype is intended to test if Q3D approach can help users to improve their spatial thinking skills when solving reasoning problems about 3D perspectives. A computer game which implements the QD-PF is intended to be developed in the near future.

Finally, we aim to carry out psychological studies that use these models and computer games in order to answer the following research questions:

- do participants train their spatial cognition skills by playing these games?
- does this Q3D/QD-FP model help participants to understand the questions?
- can these logical algorithms be applicable to more sophisticated cognitive systems (i.e. cognitive robotics)?

Acknowledgments

The support by the project *Cognitive Qualitative Descriptions and Applications* (CogQDA) funded by the U. Bremen through the *04-Independent Projects for Postdocs action*.

[3] SWI-Prolog: http://www.swi-prolog.org/

References

1. Anguera, J.A., Boccanfuso, J., Rintoul, J.L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., Gazzaley, A.: Video game training enhances cognitive control in older adults. Nature 501(7465), 97–101 (09 2013)
2. Carroll, J.: Human Cognitive Abilities: A Survey of Factor-Analytic Studies. Cambridge University Press (1993)
3. Ekstrom, R.B., French, J.W., Harman, H.H., Dermen, D.: Manual for Kit of Factor-Referenced Cognitive Tests. Princeton, NJ (1976)
4. Falomir, Z.: Towards a qualitative descriptor for paper folding reasoning. In: Proc. of the 29th International Workshop on Qualitative Reasoning, New York, USA (2016)
5. Falomir, Z.: A qualitative model for reasoning about 3D objects using depth and different perspectives. In: Lechowski, T., Waga, P., Zawidzki, M. (eds.) LQMR 2015 Workshop. Annals of Computer Science and Information Systems, vol. 7, pp. 3–11. PTI (2015)
6. Falomir, Z., Oliver, E.: Q3d-game: A tool for training users 3d spatial skills. In: Wang, M., Callaghan, V., Augusto, J.C. (eds.) Symposium on Future Intelligent Educational Environments and Learning, SOFIEE. London, UK (2016)
7. Falomir, Z., Oliver, E.: Towards improving users 3d spatial skills using a qualitative 3d descriptor and a computer game. In: Barkowsky, T., Falomir, Z., Schultheis, H., van de Ven, J. (eds.) 13th Biannual Conference of the German Society for Cognitive Science (in these proceedings). Bremen, Germany (2016), poster
8. Lloyd, J.W.: Foundations of logic programming. Symbolic computation: Artificial intelligence. Springer-Verlag, 2nd, extended edition (1987)
9. Lovett, A., Dehghani, M., Forbus, K.: Learning of qualitative descriptions for sketch recognition. In: Proc. 20th Int. Workshop on Qualitative Reasoning (QR), Hanover, USA, July (2006)
10. Lovett, A., Forbus, K.: Cultural commonalities and differences in spatial problem-solving: A computational analysis. Cognition 121(2), 281 – 287 (2011), http://dx.doi.org/10.1016/j.cognition.2011.06.012
11. Lovett, A.M., Schultheis, H.: Modeling spatial abstraction during mental rotation. In: Bello, P., Guarini, M., McShane, M., Scassellati, B. (eds.) Proceedings of the 36th annual conference of the Cognitive Science Society. pp. 886 – 891. Cognitive Science Society (2014)
12. Newcombe, N.: Picture this: Increasing math and science learning by improving spatial thinking. American Educator 34(2), 29–35 (2010)
13. Palmer, S., Rosch, E., Chase, P.: Canonical perspective and the perception of objects. Attention and Performance IX pp. 135–151 (1981)
14. Sinclair, N., Bruce, C.D.: Spatial reasoning for young learners, Research Forum. In: Proc. 38th Conf. Int. Group for the Psychology of Mathematics Education and 36th Conf. North American Chapter of the Psychology of Mathematics Education (PME 38 / PME-NA 36). pp. 173–205. Vancouver, Canada (July 2014)
15. Sorby, S.A.: Educational research in developing 3d spatial skills for engineering students. International Journal of Science Education 31(3), 459–480 (2009)
16. Spence, I., Feng, J.: Video games and spatial cognition. Review of General Psychology 14(2), 92 (2010)
17. Sterling, L., Shapiro, E.: The Art of Prolog (2nd Ed.): Advanced Programming Techniques. MIT Press, Cambridge, MA, USA (1994)
18. Wai, J., Lubinksi, D., Benbow, C.P.: Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. Journal of Educational Psychology 101(4), 817–835 (2009)
19. Yue, J.: Spatial visualization by realistic 3D views. Engineering Design Graphics Journal 72(1), 28–38 (2008)

Detecting and Discouraging Non-Cooperative Behavior in Online Rating Tasks

Jana Häussler and Tom Juzek

More and more researchers use crowdsourced rating tasks for data collection. In a rating task, participants are asked to evaluate some stimulus with respect to a given scale (e.g. they evaluate the similarity of two stimuli on a 7-point scale). However, previous studies demonstrated that crowdsourcing is quite susceptible to non-cooperative behavior (NCB), i.e. some participants are not complying with the task. Critically, NCB has a significant impact on the quality of the results that goes beyond mere noise.

This workshop presents response-time based strategies for detecting and discouraging NCB. In Session 1, we motivate their relevance, outline their functioning, and walk through the statistical part. We will show why a median-based criterion is more effective than a mean-based or absolute one and we will justify a response-time-based warning mechanism that discourages NCB effectively. Common platforms used for crowdsourced ratings tasks, e.g., Amazon Mechanical Turk or Prolific Academic, do not offer response times, let alone real-time access to them. Session 2 therefore provides the hands-on knowledge necessary for setting up an external rating website that allows the researcher to collect response times, to fully randomise items (with a Fisher-Yates shuffle), to implement the on-line/real-time warning mechanism, to intersperse booby trap items, and to collect personal data from the participants (using JavaScript and PHP; the code will be provided and explained, no prior technical knowledge is required).

Kant and Cognitive Science

Tobias Schlicht

Theoretical positions from historical figures in philosophy are not only interesting in their own right but can sometimes be especially helpful in teaching us systematic ways of inquiry that are ignored or simply unknown in contemporary debates. It has been claimed that many of Kant's ideas make him the intellectual godfather of cognitive science (e.g. his distinction of percepts and concepts, his method of transcendental argument). In several recent publications, authors have suggested that various claims from Kant's tentative Philosophy of Mind not only have counterparts in the contemporary cognitive science of the mind but can guide cognitive science in its quest to discover the function and nature of consciousness, perception and other phenomena. This tutorial has two purposes: First, to (a) outline central claims of Kant's philosophy of Mind. This is no easy task since Kant has not fully developed a full-fledged theory of consciousness or mental phenomena; rather, everything he has to say about the structure and function of mental phenomena is in the service of his epistemological project of developing a theory of knowledge. The second purpose is to (b) situate Kant's claims in contemporary debates on consciousness, (c) to evaluate which of his claims are still of use for a thoroughly naturalist approach to the mind and, more specifically (d) to evaluate whether recent claims that recent developments in cognitive neuroscience suggest a "Kantian brain" are justified.

Workshop on Creativity

Bipin Indurkhya

Human creativity has always fascinated psychologists and cognitive scientists. In the last fifty years or so, many cognitive aspects of creativity have been studied, and based on them many techniques for stimulating creativity have been developed. In this workshop, you will participate in a creativity-stimulating exercise that is based on one such technique. There are no prerequisites for participating, except to bring a fresh and open mind. This workshop is related to my talk in the KogWis 2016 symposium PROSOCRATES: Problem Solving, Creativity and Spatial Reasoning in Cognitive Systems.

Introduction to Cognitive Modeling with ACT-R

Nele Rußwinkel, Sabine Prezenski, Marc Halbrügge, and Stefan Lindner

ACT-R is the implementation of a unified theory of human cognition. It has a very active and diverse community that uses the architecture to model laboratory tasks as well as applied scenarios. The structure of ACT-R is oriented on the organization of the brain. This cognitive architectures states to be hybrid since it holds symbolic and subsymbolic components. The aim of working on cognitive models with a cognitive architecture is to understand how bottlenecks and errors occur in human behaviour occur.

In this tutorial the cognitive architecture ACT-R is introduced (J. R. Anderson, 2007: How can the human mind occur in the physical universe? New York: Oxford University Press). The focus of the tutorial is on the symbolic parts. In the beginning a short overview about recent work and ACT-R's benefit for applied cognitive science is given. Then a short introduction of the background, structure and scope of ACT-R is provided. Two hands-on examples of how to write ACT-R models are the core part of the tutorial. The first short example introduces important mechanisms of ACT-R (productions and chunks). This is followed by an in-depth introduction on mechanisms such as visual and manual processing. For the second example, the participants work on their own model version of a letter-selection task. Assistance and advice will be given during the exercises. Different solutions for the second example will be discussed. In the end information on further mechanisms of ACT-R such as subsymbolic components for learning processes are given.

No prior experience or programming knowledge is required. Please bring a laptop and preferably download the ACT-R software (stand alone version) prior to the event (http://act-r.psy.cmu.edu/software/).

Bayesian Data Analysis: Main Ideas, Practices, and Tools

Michael Franke and Fabian Dablander

Bayesian approaches to statistical inference are often portrayed as the new cool kid in town and heralded as superior to classical techniques. Naturally, the hype is also perceived critically. This course is meant to critically introduce the Bayesian approach in a nutshell. Participants who are as of yet unfamiliar with it will receive enough information to form an opinion and to know where to obtain more information that suits their needs. Those who are familiar with the main ideas can benefit from a concise rundown of the most important recent developments. In particular, this course will do two things: (i) on the conceptual level, we provide an overview of the main ideas, advantages, and challenges of Bayesian data analysis, in direct comparison to classical approaches; (ii) on a practical level, we give an executive summary of some of the most recent and convenient tools for hands-on Bayesian data analysis.

Representational Dynamics of Problem Solving in Imagery: An Exploratory Case Study

Benjamin Angerer[1], Cornell Schreiber[2], and Stefan Schneider[3]

[1] MEi:CogSci Programme, University of Vienna, Austria
bangerer@uni-osnabrueck.de
[2] Department of Philosophy, Research Platform Cognitive Science, University of Vienna, Austria
cornell.schreiber@univie.ac.at
[3] Institute of Cognitive Science, University of Osnabrück, Germany
stefan.schneider@uni-osnabrueck.de

Abstract In order to investigate the individual dynamics of representational change in problem solving, we developed a complex imagery manipulation and problem solving task (iterated paper folding) and conducted several single-case studies. Here, we present a cognitive task analysis based on a trace analysis of the gathered verbal protocols, concentrating on the varieties of representations and procedures subjects develop, as well as the circumstances of the subjects' learning events.

Keywords: problem solving, spatial imagery, representational change, cognitive task analysis, single-case studies

From everyday experience we know that both, the initial construal of a problem and the ongoing search for better representations can be non-trivial processes. Yet, for reasons of experimental methodology many tasks employed in problem solving research are very easily understood by the subjects, rendering it possible to analyse behaviour in terms of prescribed or normative problem spaces.

However, by the same token the straight-forward structure of such problems pre-empts subjects having to find adequate representations for themselves and refine them. Consequently, we still do not have a detailed account of how task and problem representations come about and change [2].

This blind spot is aggravated when we consider that problem solving might take place in multiple problem spaces at once [4,1], with complex interactions between them, and each of which being subject to change over time [5, p.9].

To allow investigating how representations develop and change over time, we developed a complex imagery manipulation and problem solving task in which subjects are asked to repeatedly mentally cross-fold sheets of paper (up to five times) and at later stages to predict the resulting sheet geometry (cf. Fig. 1). Specifically, the task was designed with the following properties in mind:

(1) It does not require substantial prior domain knowledge, (2) it is sufficiently difficult to remain so for several sessions, (3) its difficulty can scale easily for follow-up tasks (increasing the number of folds), (4) the task instruction is deliberately

under-specified (allowing subjects to develop their own task representation and varieties of folding procedures), (5) it nevertheless has well-defined solutions, and (6) through its difficulty, duration and under-specification it gives ample opportunity to develop new and better fitting representations (ranging anywhere from action-based (motoric and kinaesthetic) over sheet-based (3D deformations) and figural (constellations of edges) to more abstract, symbolic forms). Since this task specifically aims to induce a multiplicity of representations, its study calls for an open, qualitative methodology with a minimum of pre-imposed structure.

We present an extensive cognitive task analysis of our task (cf. also [3]), based on trace analyses of verbal protocols from 5 subjects, gathered with a carefully designed introspection method. The analysis will particularly pay attention to the variety of representations and procedures subjects construct, as well as to the respective prerequisites of these developments.

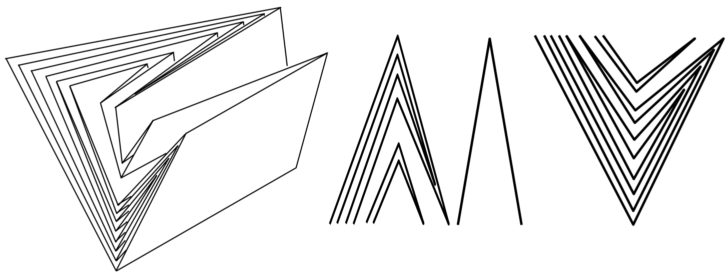


Figure 1. A 3D illustration of the 5th right-handed cross-fold and 2D illustrations of both of its complex sides

References

1. Burns, B.D., Vollmeyer, R.: Problem Solving: Phenomena in Search of a Thesis. In: Gleitman, L.R., Joshi, A.K. (eds.) Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society. University of Philadelphia, PA (2000)
2. Ohlsson, S.: The Problems with Problem Solving: Reflections on the Rise, Current Status, and Possible Future of a Cognitive Research Paradigm. The Journal of Problem Solving 5(1), 101–128 (2012)
3. Schreiber, C.: Topologie | Konstruktion einer Kreuzfaltung: Selbstbeobachtungen zum lebendigen Denken. In: Raab, T., Eder, T. (eds.) Selbstbeobachtung: Oswald Wieners Denkpsychologie, pp. 191–216. Suhrkamp, Berlin (2015)
4. Schunn, C.D., Klahr, D.: The Problem of Problem Spaces: When and How to Go beyond a 2-Space Model of Scientific Discovery. In: Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society: July 12-15, 1996, University of California, San Diego. vol. 18, pp. 25–26. Erlbaum, Mahwah, NJ (1996)
5. VanLehn, K.: Problem Solving and Cognitive Skill Acquisition. Tech report – AIP 32, Carnegie Mellon University (1988)

Twin compatibilities: Studying spatial cognition with social Simon stimuli

Pamela Baess & Christina Bermeitinger

University of Hildesheim, Department of Psychology, Germany

baessp@uni-hildesheim.de

Abstract
Faster responses when stimulus and response are compatible have been numerously shown with the so-called Simon-effect. However, previous studies mostly used stimuli embedding only one stimulus-response compatibility at the time. In our new experiment, we used stimuli containing two different kinds of stimulus-response-compatibilities at the same time and investigated whether the two different kinds of task-irrelevant spatial information would be processed differently.

Keywords. Simon effect; allocentric reference frame; egocentric reference frame

1 Introduction

Spatial information, even though completely task-irrelevant, is hard to be ignored. This has consistently been reported in the so-called Simon-Task (for a review, Simon, 1990). In this task, participants are asked to respond to a certain stimulus feature (e.g. color) by pressing one of two response buttons (left vs. right) while ignoring the other stimulus feature (e.g. the position on the screen). Although the target position is irrelevant for the relevant task, participant's performance turns out to be faster when the task-irrelevant stimulus feature is congruent with the side of the response buttons than when it is incompatible. Thus, this kind of stimulus-response-compatibility (SRC) effects illustrate that the elements of the stimulus and the response sets interact. Previous studies mainly used abstract or simple figures as the task-relevant stimulus material addressing various questions along the compatibility between the stimulus and the response. There are only few studies using more social stimuli such us hands or eyes (Lameira, Pereira, Fraga-Filho, & Gawryszewski, 2015; Pomianowska, Germeys, Verfaillie, & Newell, 2011; Ricciardelli, Bonfiglioli, Iani, Rubichi, & Nicoletti, 2007). Moreover, the target stimuli used in the Simon Task have almost exclusively been presented in an egocentric perspective, i.e. relative to the mid-line of the screen) and containing only one stimulus-response compatibility. In the present study, we used the schematic drawing of a human, i.e. stick-figure manikins, as stimulus material. The manikin carried a differently colored ball in either his left or right hand, thus, enabling two different stimulus-response compatibilities (regarding both, the manikin position on the screen and the ball position) to occur. Manipulating the reference frame, we presented the manikin in an egocentric perspective (one manikin on the screen) vs. and allocentric perspective (9 identical manikins on the screen). This newly developed paradigm promotes studying spatial cognition in a more real life scenario using different reference frames and multiple SRCs. Here, we will present the results of two different experiments using this paradigm.

2 Methods

117 subjects participated in this Experiment. The participants were randomly assigned to one of two between-subject groups (1 manikin display vs. 9 manikin display). Further, the color of the ball (yellow vs. blue; corresponds to the response buttons) and the position of the manikin on the screen (left vs. right) was manipulated. For more details on the experimental procedure, see Figure 1.

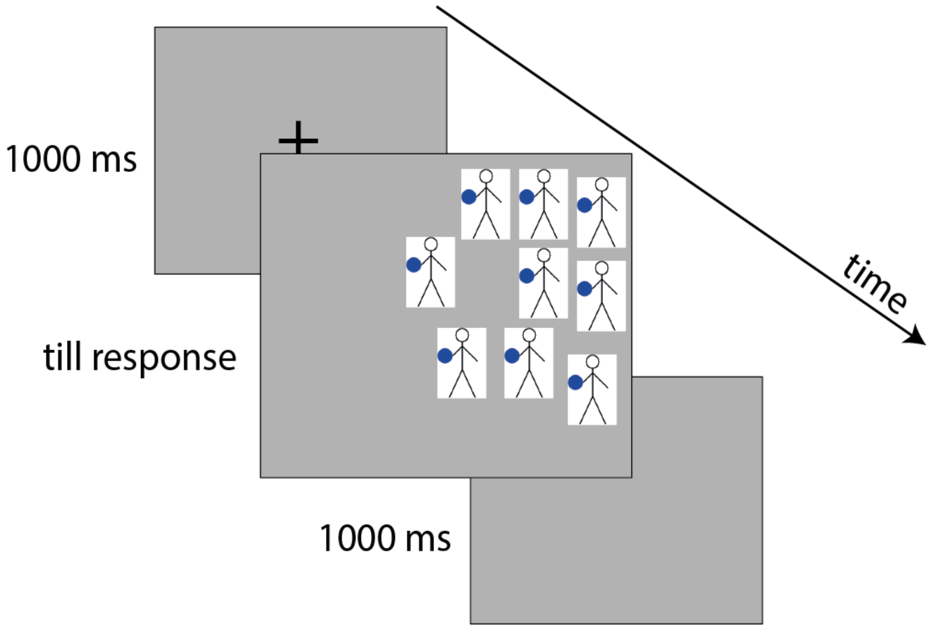
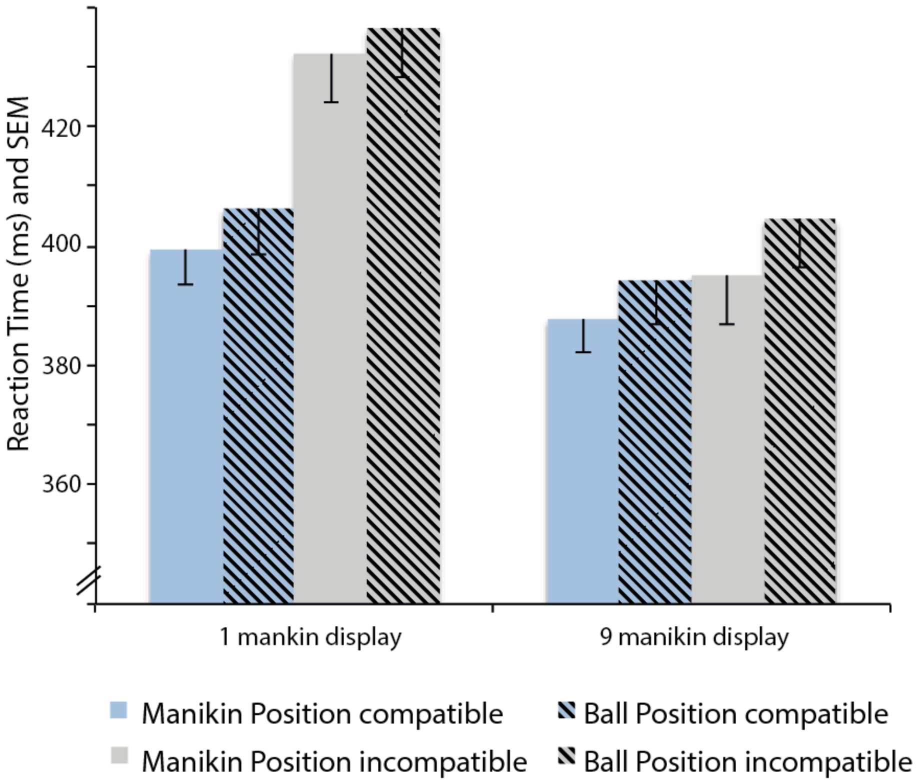


Fig. 1. Trial structure of the experiment. The position of the manikin was randomly drawn out of 18 possible positions under the constraint to have more manikins on one side (left vs. right). For the 1-manikin display group, the only one manikin was presented on the screen.

3 Results and Discussion

Even though the spatial position of the ball or the position of manikin was completely task-irrelevant, both spatial information have been processed further (Figure 2). Fastest responses were yielded when both kind of spatial information were compatible, i.e. the ball side matched the presentation side of the manikin on the screen. However, the size of the SRCs was differently affected by the more egocentric vs. allocentric reference frame. Particularly, the Simon Effect due to the position of the manikin decreased drastically under the allocentric compared to the egocentric reference frame.

Fig. 2. Reaction times and SEM, separated for the egocentric vs. allocentric stick-figure display.



4 References

Lameira, A. P., Pereira, A., Fraga-Filho, R. S., & Gawryszewski, L. G. (2015). Stimulus-response compatibility with body parts: a study with hands. *Exp Brain Res, 233*(7), 2127-2132.

Pomianowska, I., Germeys, F., Verfaillie, K., & Newell, F. N. (2011). The role of social cues in the deployment of spatial attention: head-body relationships automatically activate directional spatial codes in a Simon task. *Front Integr Neurosci, 6*, 4.

Ricciardelli, P., Bonfiglioli, C., Iani, C., Rubichi, S., & Nicoletti, R. (2007). Spatial coding and central patterns: Is there something special about the eyes? *Canadian Journal of Experimental Psychology, 61*(2), 79-90.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G.

Reeve (Eds.), *Stimulus-response compatibility: an integrated perspective* (pp. 31-86). Amsterdam: North-Holland.

The impact of sleep on navigation and consolidation of survey knowledge

Matthias-Philipp Baumann[1], Wiebke Schick[2]; Hanspeter A. Mallot[2]
Cognitive Neuroscience, Dept of Biology, University of Tübingen
* Corresponding author: matthias-philipp.baumann@student.uni-tuebingen.de

Background

Spatial cognition involves complex processes that include the integration of different sources of information. For navigation, beneficial effects of sleep have been reported (e.g. Peigneux et al, 2004; Wamsley, 2010). Here, we investigate the influence of sleep on memory consolidation in the process of learning a new environment and establishing a general, functional representation of it. The experimental environment was an iterated hexagonal y-maze in a VR design that had to be navigated in two sessions, with or without an intermittent sleep phase (Fig. 1; Gillner&Mallot, 1998).

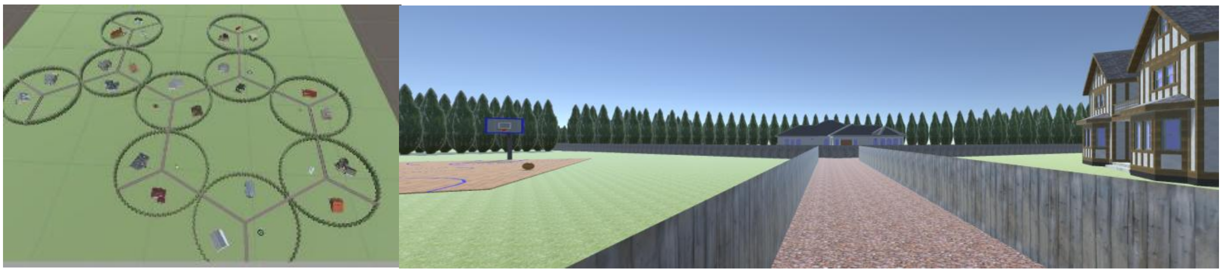


Fig. 1: Left: Survey of the experimental environment. Right: View of an intersection

Methods

The virtual environment was presented with the virtual reality headset Oculus Rift 2. It contained 10 places, each of them characterized by 2 houses and another typical inner-city object, e.g. a bus stop. Participants performed two sessions of navigational tasks, namely the Travelling-Salesman-Problem. Sleep was controlled for by dividing the subjects into two groups that performed the sessions in different orders: One group performed the first session in the evening, the second in the morning of the following day, and slept at home in between the two sessions, the other group was tested in the morning and in the evening of the same day. In the second session, subjects had to navigate routes that allowed for the shortest way by combining segments of the first session in new ways. After the second VR-session, subjects had to answer a questionnaire and draw a sketch of the environment.

Participants

22 subjects were assigned to one of the two testing groups. Each group consisted of 11 subjects (6 male, 5 female).b

Variables of interest

The navigational performance was evaluated through the errors, the error rate and the time required for completion of the navigational tasks. An error was defined as passing a place that increased the distance to the next target. The error rate was the ratio between wrong and right decisions. The sketch quality was measured by predefined criteria (e.g. the recall of the general layout, the number of correctly recalled places).

Results

In both conditions, we found a significant improvement of the required time and the error rate (time: $p < 0.05$, effect size: $\eta2= .813$; error rate: $p < 0.05$, effect size: $\eta2= .539$; Fig. 2).

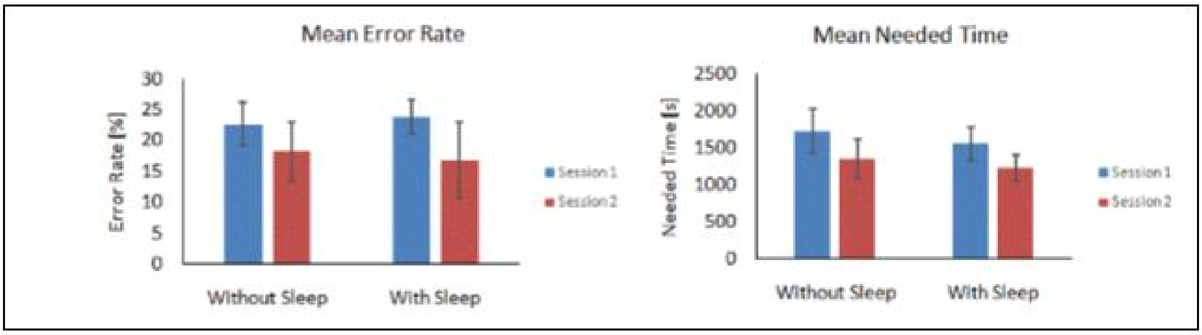


Fig. 2: Mean navigation errors per session.

The quality of the sketches of the experimental environment after the second session correlated both to the required time and the navigational errors (r [20]= - 0.575; p= 0.005; Fig. 3).

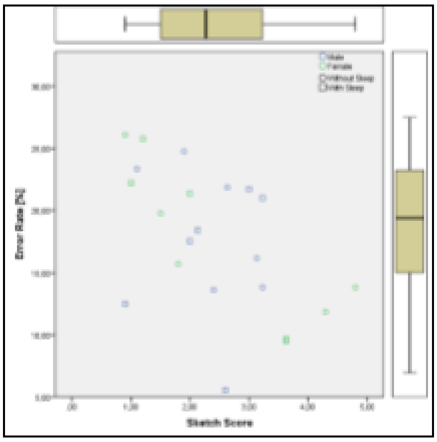


Fig. 3: Correlation between the navigation errors and the quality of the sketches.

Conclusions

1. A lower error rate and faster task completion was observed in the second session in both conditions, which indicates learning.
2. There was a correlation between the navigational errors and quality of the sketch map.
3. A significant improving effect of sleep was not found.

Due to the design of the experiment, it is not possible to deferr whether these findings indicate no beneficial effect of sleep on the consolidation of survey knowledge - or whether there is a general lower quality of learning in the evening. More data is needed for a better differentiation between all factors influencing learning and memorization processes.

References

Gillner, S., Mallot, H.A. (1998). *Navigation and acquisition of spatial knowledge in a virtual maze*. Journal of Cognitive Neuroscience 10 445-463.

Peigneux P, Laureys S, Fuchs S, Collette F, Perrin F, Reggers J, Phillips C, Degueldre C, Del Fiore G, Aerts J, et al. 2004. *Are spatial memories strengthened in the human hippocampus during slow wave sleep?* Neuron 44: 535–545.

Wamsley, E. J., Tucker, M. A., Payne, J. D., &Stickgold, R. (2010). *A brief nap is beneficial for human route learning: The role of navigation experience and EEG spectral power*. Learning & Memory, 17(7), 332-336.

Moral Decision Making in Autonomous Vehicles

An Empirical Behavioral Approach Suggesting a Modified Utilitarianism

Lasse T. Bergmann*, Silja Timm, Max Wächter, Anke Dittmer, Felix Blind, Carmen Meixner, Larissa Schlicht, Anja Faulhaber, Juhee Jang, Aalia Nosheen, Simeon Kraev, Max Räuker, Leon Sütfeld, Achim Stephan, Peter König, Gordon Pipa

*direct contact requests to: lbergman@uos.de
Institute of Cognitive Science
University of Osnabrück

Abstract. In order to investigate the commonsense morality guiding human behavior, we developed an experimental paradigm to test critical situations in traffic. The experimental data allows inferences towards the moral principles guiding human behavior in street traffic. These insights will be helpful in creating norms for autonomous vehicles to follow.

Keywords: Autonomous Vehicles, Moral Decisions, Utilitarianism

Autonomous vehicles are the emerging technology of the next decade. Like smartphones, autonomous vehicles will become ubiquitous very quickly. Increased safety and comfort, as well as economic benefits, will drive the development and hasten adoption.
But, as with any technology, many issues will arise, since autonomous driving will affect a vast number of humans. Furthermore it will require interaction between humans and machines on a scale and with stakes higher then ever before. One of the substantial problems will be moral decision making. Even with autonomous vehicles accidents can't be completely precluded. Hence the situation may arise that a autonomous vehicle has to make moral decisions on its own.[4]
This problem somewhat resembles the infamous *trolley problem*. Suppose a autonomous vehicle drives along a narrow road, a child steps on the street from an unobserved spot so that the vehicle can only avoid the collision by driving on the sidewalk, but there is not enough time to brake. On the sidewalk however another person would be hit. The autonomous vehicle would need to make a moral decision, which person should be hit.[2]
Appealing to moral theory in this situation is rather unhelpful, as it is unclear which moral theory should guide the actions. As in the trolley problem moral theories make contradictory claim as to what is the right action. But what criterion should then guide moral decision making?
In case of conflicting moral theories and a practical need to resolve these conflicts, consulting commonsense intuitions became popular recently.[3] We can learn from the judgements people make in these situations which kind of coremorality is present in humans. Furthermore this ensures that the judgements

derived from this process are acceptable to society, as people tend to accept what comes naturally to them.[1] Acceptance is an important consideration here, as a domain-specific moral algorithm for autonomous cars could only be legitimated through a societal agreement. This provides the normative justification for using these algorithms, that their neuropsychological basis lacks.

This transforms the core question of ethics, *how should one act*, in the case of autonomous vehicles into an empirical question. Through virtual reality experiments, in which subjects were put into critical moral dilemma situations in traffic, we investigated which moral principles guide human decision making in traffic. In a previous study it was shown that a behavioral virtual reality paradigm gives strong indications on commonsense morality and that behavior is a good indicator of acceptance of behavior.[5] A refined study, that was more specific to problematic situations, provided further insight into commonsense morality. Subjects tended to favor *utilitarian* decisions, but didn't commit to *act utilitarianism* strictly. The experiments suggested, that commonsense morality may adhere to additional rules. Especially situations in which utilitarianism required self-sacrifice or violation of traffic rules were investigated. Furthermore the impact of age on moral judgement, the special role of children in traffic, and consideration of vulnerability were investigated.

Results indicate that people were not totally comfortable with the utilitarian command to sacrifice oneself for the good of others, while still a majority conformed to the utilitarian principle, another portion only did so for larger groups, while a small group of participants refused to sacrifice themselves completely. Similar behavior was observed in connection with traffic rules, even though a strong majority ignored rules in favor of conformity utilitarian commands, for a small group of subjects adherence to traffic rules seemed to be more important than aggregative utilitarian norms. Furthermore a tendency to kill elderly people rather than younger adults or children was found, which may point towards subjects assigning value based on the life people still have ahead of themselves.

References

1. Jean-François Bonnefon, Azim Shariff, and Iyad Rahwan. Autonomous vehicles need experimental ethics: Are we ready for utilitarian cars? *arXiv preprint arXiv:1510.03346*, 2015.
2. Noah Goodall. Ethical decision making during automated vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board*, (2424):58–65, 2014.
3. Joshua Greene. *Moral tribes: emotion, reason and the gap between us and them.* Atlantic Books Ltd, 2014.
4. Alexander Hevelke and Julian Nida-Rümelin. Selbstfahrende autos und trolleyprobleme: Zum aufrechnen von menschenleben im falle unausweichlicher unfälle. *Jahrbuch für Wissenschaft und Ethik*, 19(1):5–24, 2015.
5. L.R. Sütfeld, R. Gast, P. König, and G. Pipa. Ethical decision: Profound impact of assessment methodology, well-described by one-dimensional value of life scale. 2016.

Decoding spatial auditory attention using ear EEG

Martin G. Bleichner, Bojana Mirkovic, Stefan Debener

Department of Psychology, Neurospsychology Lab, European Medical School, Cluster of Excellence Hearing4all, University of Oldenburg, Oldenburg, Germany

Introduction: Spatial auditory attention allows us to select auditory objects in complex (naturalistic) auditory scenes based on spatial cues. It has been shown that properties of the attended auditory source are reflected neurophysiologically as measured by magnetoencephalography (Ding and Simon, 2012) and elecetroencephalography (Mirkovic et al., 2015). Choi et al. (2013) showed that the direction of auditory attention could be predicted based on the temporal structure of the ERP in respect to the sound onsets in an attended musical segment. Here we replicate the result of Choi et al. (2013) comparing a classical EEG cap setup with a behind-the-ear electrode array (Debener et al., 2015, figure 1 A).
Methods and Methods: We simultaneously recorded 84-channel scalp EEG and 18 channel ear EEG from 20 healthy volunteers (figure 1 A). Participants were instructed to attend to the left or right stream of three concurrently presented three seconds music streams (but never to the center stream, figure 1 B). The streams differed in their direction of origin (center, left, right), musical instrument, tone pitch, and the number of tones (3, 4, and 5 respectively). For the attended stream participants had to indicate whether the pitch sequence was ascending, descending or alternating. For both setups the grand average ERP for the entire segment and for the individual tones were computed as well as the corresponding effect sizes. The single trials were classified using a leave one out template matching approach.
Results: The grand average ERP for the single tones clearly shows an attention effect on the N1-P2 complex for both setups (figure 1 C). The signal amplitude is approximately twice as large for the cap compared to the cEEGrid. The effect size (Hedges' g) over time was very similar for the two setups in its temporal evolution and magnitude. The grand average ERP for the entire segment (figure 1 E) show clear differences between the attend-left and attend-right condition, and reflect the temporal structure of the attended stream. The single trials were classified above chance level for both systems for 16 (cEEGrid) and 17 (cap) out of 20 datasets. The median accuracy was 66% (range 57% to 85%) for the cEEGrid and 70% (range 56% to 89%) for the cap. There was no significant difference between the classification accuracies (Wilcoxon signed rank test: $W=1$, $Z=-1.4777$, $p=0.145$, $r=0.33$). Instead, the accuracies were significantly correlated ($r=0.7127$, $p<0.001$) as illustrated in figure 1D.
Discussion: We could replicate the results of Choi et al. (2013) using the cap-EEG and found comparable results with the cEEGrid in respect of the morphology of the ERP, the effect size and the single trial classification. We argue that concealed behind-the-ear EEG can be an alternative for cap-based EEG in monitoring auditory attention and also be used to study the relationship between neural activity and behavior in naturalistic settings with minimal interference with the users' normal behavior.

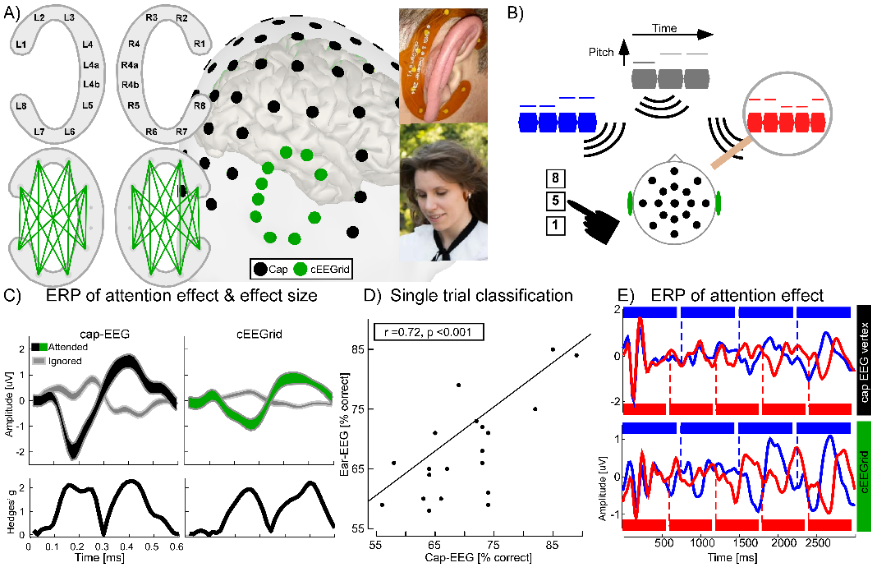


Fig. 1. A) cEEGrid electrode layout. R4a and R4b were used as ground and reference electrode. Below: the green lines indicate the bipolar channel pairs that were used for the analysis. Middle: the electrode positions of cap (black) and cEEGrid (green) shown on a standard anatomy. Right: cEEGrid electrodes attached behind the ear. B) Auditory attention paradigm. Three music streams are presented originating from the left (blue), center (gray) and right (red). The participant had to attend to one of the streams and to indicate whether the sound sequence was ascending, descending or alternating. C) Top: Grand average ERP of attended tones (green and black) and unattended tones (grey), excluding the first tone of each sequence. Below: effect size over time (Hedges' g). D) Single trial classification accuracy for cap EEG and cEEGrid are significantly correlated with each other. E) The grand average ERP for the attend-left (blue, four tones) and the attend-right (red, five tones) follow the onsets of the individual tones (dashed vertical lines indicate tone onsets, horizontal bars indicate tone onset and duration).

References

Choi, I., Rajaram, S., Varghese, L.A., Shinn-Cunningham, B.G., 2013. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. Front. Hum. Neurosci. 7.

Debener, S., Emkes, R., De Vos, M., Bleichner, M., 2015. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. Sci. Rep. 5, 16743.

Ding, N., Simon, J.Z., 2012. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 107, 78–89.

Mirkovic, B., Debener, S., Jaeger, M., Vos, M. De, 2015. Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. J. Neural Eng. 12, 046007.

Learned knowledge about the co-actor's behavior influences performance in a joint visuomotor task

Artur Czeszumski[1], Ernesto Andrés López Montecinos[1], Chiara Carrera[1], Anette Aumeistere[1], Ann Xavier[1], Basil Wahn[1,*], Peter König[1, 2,*]

[1] Institute of Cognitive Science, Universität Osnabrück, Osnabrück, Germany
[2] Institut für Neurophysiologie und Pathophysiologie, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany
[*] shared senior authorship

Keywords: Social cognition, Joint action, Visuomotor task

When humans act together, members of a dyad use prior knowledge about his co-actors as well as perception-action couplings [2,5–8]. A current view implies that direct action-perception couplings suffice for joint action performance [4,5]. Due to the limited amount of evidence to support this claim, we investigate the influence of knowing the co-actor's identity on joint performance. More precisely, within a joint visuomotor task involving two dyads simultaneously, we test the role of prior knowledge about the co-partner's identity.

The joint visuomotor task required a systematic switching of co-partner: Each participant was paired with 3 different co-partners (Figure 1). The task consisted in moving a circular cursor towards a target position on a computer screen, with the goal of reaching the target position as fast as possible, while following the specified trajectory. We converted the time needed and the cumulative deviations from the specified trajectory to a single penalty score. Participants reduced or incremented the cursor's velocity in horizontal or vertical directions by button presses. In the joint task, each individual had control only of one of the two spatial dimensions (horizontal or vertical) while the co-partner was responsible for the other dimension. Before starting the experiment, each participant chose one identification avatar that is maintained for the whole duration of the experiment (Figure 2).

In a first training phase, participants learned how to coordinate their actions with each co-actor separately by playing a sufficient amount of trials (24 trials with each co-actor, divided into randomized 6 blocks of 4 trials), in which they had access to the identity information. We have chosen these parameters based on the results of a previous pilot study. Subsequently, they performed the same task but with a random manipulation of this identity information (experimental phase). More precisely, the identity of the co-partner was either displayed (informed, 16 trials with each co-actor), missing (uninformed, 16 trials with each co-actor), or wrong (misinformed, 4 trials with each co-actor). 20 quadruples (N=80, 46 females, mean age = 22.5 years, SD = 5.7 years) participated in the experiment.

A decrease of the penalty score over blocks suggested that subjects learned how to coordinate with different partners in turns during the training phase (Figure 3). We performed linear mixed effects analysis of the relationship between penalty score and dimension of the trajectory. We entered dimension as a fixed effect, and intercepts for quadruples as a random effect into the model. Using a likelihood ratio test for model comparisons, we found that the penalty score for vertical trajectories was significantly lower than for horizontal trajectories ($\chi^2(1)=57.31$, $p <. 001$, mean difference = 0.119, SD ± 0.02). To test differences between conditions (informed, uninformed), we performed linear mixed effects analysis of the relationship between penalty score and the factors condition and dimension. As a random effect, we added intercepts as well as random slopes for the effect of condition for each quadruple. Using model comparisons, we found a trend towards significance for the factor condition ($\chi^2(1)=6.59$, $p = .065$). In particular, the penalty score was 0.042 (SD ± 0.02) lower in the uniformed condition than in the informed condition. To assess how the dissimilarity between pairs within a quadruple relates to the difference between the informed and uninformed condition, we first computed for each pair within a quadruple a feature vector comprised of all measured dependent variables (i.e., Penalty, Trial completion time, Velocity, Button press frequency, and Perspective taking score [1]) to characterize the joint performance. We defined a dissimilarity score for each quadruple by calculating the averaged Euclidean distance between feature vectors (Penalty, Time, Velocity, Button press frequency, and Perspective taking score [1]) for all pairs within the quadruple. Furthermore, we correlated the dissimilarity scores with the difference between the informed and uninformed condition and found a significant negative correlation for horizontal trajectories ($r = -0.47$, $p = .037$, Figure 4) but not for vertical trajectories ($r = -0.08$, $p = .743$, Figure 5).

Taken together, these findings suggest that it is easier to coordinate in time in a joint visuomotor task when humans do not use prior knowledge about their co-actor. However, when the dissimilarity between co-actors increases then the information about the co-actor's identity becomes more relevant.

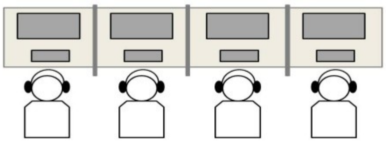


Figure 1: Experimental Set-Up

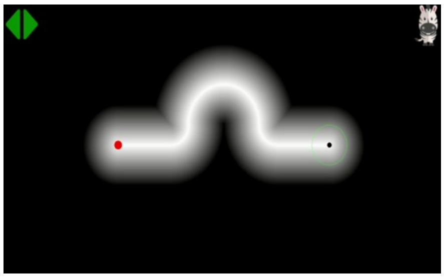


Figure 2: Experimental Design for Informed Trial

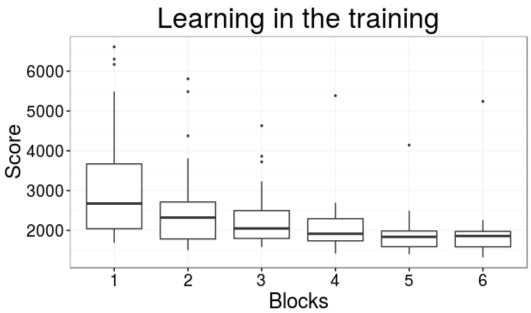


Figure 3: Changes in the overall score over the 6 block trials in the first phase (training)

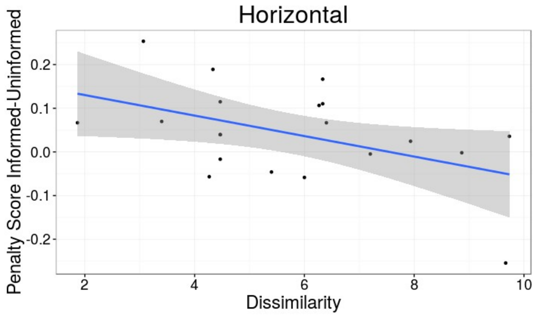


Figure 4: y-axis differences between informed and uninformed trials for each quadruple, x-axis dissimilarity values

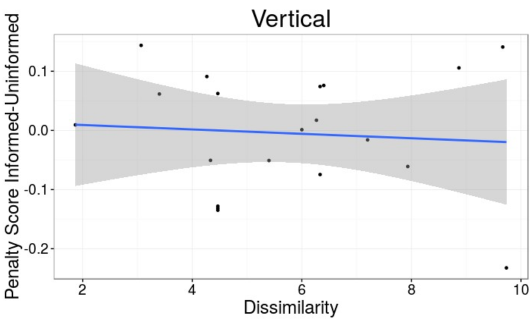


Figure 5: y-axis differences between informed and uninformed trials for each quadruple, x-axis dissimilarity values

We gratefully acknowledge the support by H2020 – H2020-FETPROACT-2014 641321 – socSMCs.

References

1. Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology, 10, 85.
2. Knoblich, G., Butterfill, S., Sebanz, N.: Psychological research on joint action: Theory and data. The Psychology of Learning and Motivation, 54, 59–101 (2011)
3. Marsh, K. L., Johnston, L., Richardson, M. J., Schmidt, R. C.: Towards radically embodied, embedded social psychology. European Journal of Social Psychology, 39, 1217–1225 (2009)
4. Schmidt, R. C., Richardson, M. J.: Dynamics of interpersonal coordination. In: Fuchs, A., Jirsa, V. K. (eds.) Coordination: neural, behavioral and social dynamics, pp. 281–308. Berlin, Heidelberg: Springer (2008)
5. Sebanz, N., Bekkering, H., Knoblich, G.: Joint action: bodies and minds moving together. Trends in Cognitive Sciences, 10(2), 70-76 (2006)
6. Sebanz, N., Knoblich, G.: Prediction in joint action: what, when, and where. Topics in Cognitive Science, 1, 353–36 (2009)
7. Sebanz, N., Shiffrar, M.: Bodily bonds: Effects of social context on ideomotor movements. In: Rossetti, Y., Kawato M., Haggard, P. (eds.) Sensorimotor foundations of higher cognition (attention and performance, XXII). pp. 267-292. Oxford, UK: Oxford University Press (2007)
8. Vesper, C., Butterfill, S., Knoblich, G., Sebanz, N.: A minimal architecture for joint action. Neural Networks, 23, 998–1003 (2010)

Quick and sustained inhibition of distractor elicited response activation in task switching

Philipp Dehmel, Kerstin Jost, Aquiles Luna-Rodriguez, Mike Wendt and Thomas Jacobsen

Philipp Dehmel, M.Sc.
Helmut Schmidt University / University of the Federal Armed Forces Hamburg
Faculty of Humanities and Social Sciences
Experimental Psychology Unit
Holstenhofweg 85
22043 Hamburg

Philipp.dehmel@hsu-hh.de

Abstract. Task performance is often affected by the presence of task-irrelevant stimuli (i.e., distractors). Although in most cases responses are facilitated if the target and the distractor of a current trial are associated with the same response (i.e., congruent condition) compared to when they are associated with different responses (i.e., incongruent condition), reversed congruency effects (i.e., better performance in incongruent than in congruent trials) have also been observed. Reversal of the congruency effect has been attributed to inhibition of distractor-elicited response activation.

In a series of experiments, we investigated the time course of distractor-target congruency effects in a task switching paradigm, involving three tasks that differed regarding perceptual dimensions (i.e., color, shape, and direction). The same task was never repeated across consecutive trials, therefore only task switch trials were administered. Target stimuli were preceded by distractors from either the current task or a different task at varying intervals. Assuming inhibition of response activation we expected the congruency effect evoked by distractors to decrease and eventually reverse with an increasing distractor-target interval.

In fact, both same- and different-task distractors yielded reversed congruency effects already in the shortest distractor-target SOA (150 ms or 250 ms, depending on the experiment). This reversal of the congruency effect was observed for all distractor-target SOAs administered (i.e., up to 1500 ms), suggesting consistent inhibition of distractor-elicited response activation in task switching situations. This inhibition is applied relatively quickly and seems to be sustained for long intervals.

The Role of Facial Mimicry in Cognitive and Emotional Empathy and Effects of Autistic Traits

Hanna Drimalla[1] and Isabel Dziobek[1]

[1]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany
hanna.drimalla@hu-berlin.de

It is well known that people automatically mimic facial expressions of others [1]. While some authors claim that mimicry helps people to recognize emotions [2], others argue that mimicry fosters not recognition but affiliation, as people mimic others more if they want to affiliate with them [3]. These two possible motivators of mimicry can be conceptualized as the two facets of empathy: cognitive empathy (recognition of mental states) and emotional empathy (feeling with another person), which dissociate in individuals with autistic conditions [4].

To directly compare the role of mimicry in these two facets of empathy and its relationship with autistic traits, we have conducted a study with 40 healthy male participants between 18 and 35 as autistic traits are more prevalent in males [5]. We measured cognitive and emotional empathy simultaneously via the condensed and revised version of the Multifaceted Empathy Test (MET) [4]. In this validated test photographs of persons in emotionally charged situations are presented. The participant is asked to identify the emotion (cognitive empathy) of the people on the pictures as well as to indicate how much she or he feels for (emotional empathy) the individuals. While the participants performed the MET, the mimicry reaction was measured via electromyography of the zygomaticus major (ZYG, “smiling muscle”) and the corrugator (COR, “frowning muscle”). A stronger reaction of ZYG (compared to COR) to positive pictures as well as a stronger reaction of COR (compared to the ZYG) to negative pictures was seen as mimicry response. Its strength was calculated as a difference score between the two muscles (ZYG minus COR). The activity was averaged from 500ms to 4000ms after picture onset to exclude any orientation response. Then the mean signal of the 500ms before the presentation was subtracted (intra-individually and trial-based) to control for baseline activity.

We found a clear pattern of mimicry for the positive and negative emotions of the MET (fig. 1). The mimicry pattern emerged during the emotional as well as the cognitive empathy pictures. The stronger the mimicry of positive emotion was the better the participants performed on the recognition of these emotions ($r(38)=0.3523$, $p<0.05$, fig. 2). Moreover, the mimicry for negative emotions correlated negatively (for cognitive items: $r(38)=-0.41$, $p<0.01$ and emotional items: $r(38)=-0.321$, $p<0.05$) with autism traits measured by a short-version of the Autism-Spectrum Quotient (AQ) [6]. Our findings support the view on mimicry as a facilitator of cognitive empathy as the participants recognized as more positive emotions as more they mimicked. In contrast, emotional empathy was associated with mimicry only on a trait(AQ) level.

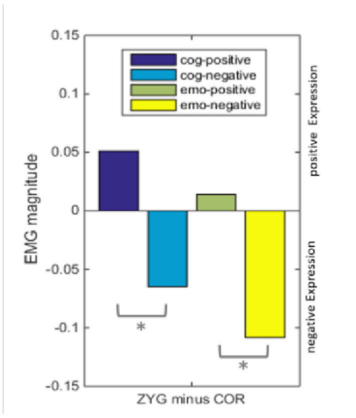


Fig. 1. Mimicry in the MET: The EMG magnitude is the difference of ZYG minus COR activity averaged over the time window from 500ms to 4000ms after stimulus onset (z-standardized and baseline-controlled)

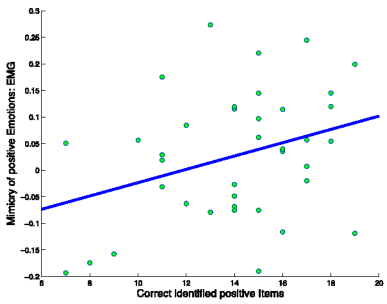


Fig. 2. Mimicry Response and Emotion Recognition in the MET (only positive emotions): The Mimicry is measured as a difference of ZYG minus COR activity, averaged over the time from 500ms to 4000ms after stimulus onset (z-standardized and baseline-controlled)

References

1. Dimberg, U., Thunberg, M., Elmehed, K.: Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*: 86-89 (2000)
2. Blairy, S., Herrera, P., Hess, U.: Mimicry and the judgment of emotional facial expressions. *Journal of Nonverbal behavior* 23, 5-41 (1999)
3. Likowski, K.U., Muhlberger, A., Seibt, B., Pauli, P., Weyers, P.: Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology* 44: 1065–1072 (2008)
4. Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A.: Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*(3), 464-473 (2008)
5. Baron-Cohen, S.: The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6): 248-254 (2002)
6. Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S.: The Autism-Spectrum Quotient (AQ)—adolescent version. *Journal of Autism and Developmental Disorders*, *36*, 343–350 (2006)

Towards Improving Users' 3D Spatial Skills using a Qualitative 3D Descriptor and a Computer Game*

Zoe Falomir[1] and Eric Oliver[1,2]

[1]Spatial Cognition Centre, Universität Bremen, Germany
[2]Universitat Jaume I, Castellón, Spain
zfalomir@informatik.uni-bremen.de, ericoliverobiol@gmail.com

1 Extended Abstract

Spatial cognition studies showed that there is a strong link between success in Science, Technology, Engineering and Math (STEM) disciplines and spatial abilities [4]. These abilities are basic for job skills like visualizing the result of a surgery, designing bridges, aircrafts; interpreting charts, maps, engineering drawings, etc. Moreover, it appears that 3D spatial skills can be developed through practice since research showed [6] that students who attended a course at university to improve their ability to visualize in 3D, improved their success and retention significantly, particularly female students.

Qualitative models that try to solve spatial cognition problems have appeared in the literature [2, 1, 3]. A qualitative descriptor for solving paper folding tests was defined [2] by establishing a correspondence between the possible folding actions and the areas in the paper where a hole can be punched. A logic-based formalization of the Fishermans Folly puzzle was proposed using qualitative spatial reasoning about strings and holes and reasoning about actions and change on these objects [1]. A qualitative model for describing 3D objects (Q3D) using depth and different perspectives [3] was defined based on designs of pieces described using 3 canonical views (top, lateral and front) since, in experimental psychology, there is support for the general idea that human object recognition involves view-dependent representations, that is, people prefer to imagine, view, or photograph objects from certain *canonical* views [5]. The Q3D approach was motivated by the fact that: (i) the German Academic Foundation uses consistent view/projection of a 3D object corresponding to a technological drawing to measure intelligence in students[1]; and that (ii) the Dental Admission Testing Program[2] by the American Dental Association includes a Perceptual Ability Test (PAT) which includes some parts on 3D object perspective reasoning. The Q3D approach allows reasoning through logics defined to test the consistency of descriptions and it can infer features of the unknown perspectives (i.e. back, down and the other lateral).

As research has shown that video game training enhances cognitive control [7] and that realistic 3D views enhanced users' performance on spatial visualization tests

* The project *Cognitive Qualitative Descriptions and Applications* (CogQDA) funded by the Universität Bremen, the European Erasmus+ Intership program and the support by the *Gesellschaft für Kognitionwissensaft* (GK) are acknowledged.

[1] *Test der Studienstiftung*: http://www.spiegel.de/quiztool/quiztool-49771.html

[2] *Dental Admission Testing Program example*: http://www.ada.org/

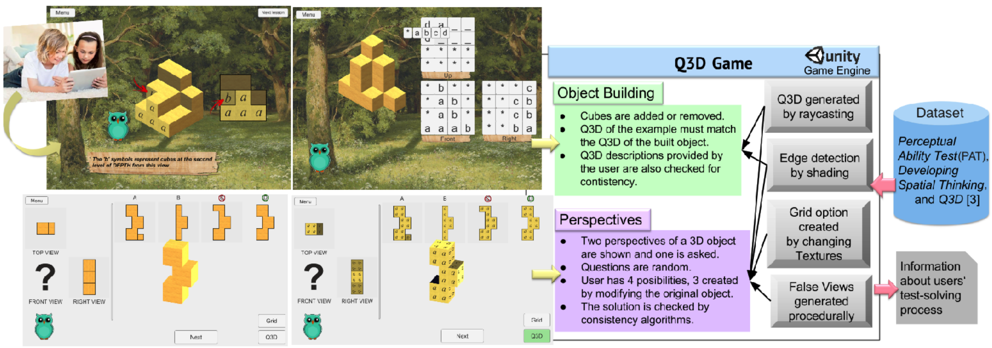


Fig. 1. Q3D-Game options and the main implementation procedures.

[8], the Q3D model was implemented in a video game, Q3D-Game. The Q3D-Game showed here is intended to test if the Q3D model can help users to improve their spatial thinking skills when solving reasoning problems about 3D perspectives. The Q3D-Game was developed using Unity 5 game engine because it provides multiplatform exporting to mobiles and tablets. The Q3D-Game runs on any tablet/mobile with an operative system Android 2.3.1 'Gingerbread' (API level 9) or upper version.

The Q3D-Game has three different playing options: (i) object building, where players learn how to build objects using cubes; (ii) object description, where players are shown how to describe an object from its perspectives using Q3D; and (iii) perspective question answering, where players are shown a dataset of questions regarding 3D perspectives built from real questions included in the PAT. Experiments on participants are carried out at the moment and results will be published in the near future.

References

1. Cabalar, P., Santos, P.E.: Formalising the fisherman's folly puzzle. Artif. Intell. 175(1), 346–377 (2011)
2. Falomir, Z.: Towards a qualitative descriptor for paper folding reasoning. In: Proc. of the 29th International Workshop on Qualitative Reasoning (2016)
3. Falomir, Z.: A qualitative model for reasoning about 3D objects using depth and different perspectives. In: Lechowski, T., Waga, P., Zawidzki, M. (eds.) LQMR 2015 Workshop. Annals of Computer Science and Information Systems, vol. 7, pp. 3–11. PTI (2015)
4. Newcombe, N.: Picture this: Increasing math and science learning by improving spatial thinking. American Educator 34(2), 29–35 (2010)
5. Palmer, S., Rosch, E., Chase, P.: Canonical perspective and the perception of objects. Attention and Performance IX pp. 135–151 (1981)
6. Sorby, S.A.: Educational research in developing 3d spatial skills for engineering students. International Journal of Science Education 31(3), 459–480 (2009)
7. Spence, I., Feng, J.: Video games and spatial cognition. Review of General Psychology 14(2), 92 (2010)
8. Yue, J.: Spatial visualization by realistic 3D views. Engineering Design Graphics Journal 72(1), 28–38 (2008)

The Cortical Network of Usability Evaluations for Unknown Tools

Mareike Gann and Marc Himmelbach

Division of Neuropsychology, Centre for Neurology,
Hertie-Institute for Clinical Brain Research
{mareike.gann@student.uni-tuebingen.de, marc.himmelbach@uni-tuebingen.de}

Abstract. In daily life we are surrounded by different tools. Most often, we encounter familiar tools, and we can retrieve information about them from our semantic knowledge. Sometimes, however, we are confronted with new and unfamiliar tools. We must analyze and interpret their properties and infer a possible use. We assumed that such sensory and cognitive processing of unfamiliar tools will place a higher demand on a previously reported left-hemispheric network comprising the IFG/vPMC, anterior and dorsal IPL, and posterior ITG/IOG. In contrast, retrieving the well-known functionality of highly familiar tools should result in higher demands on the posterior MTG which has been associated with the retrieval of information from semantic networks before.

We asked 25 healthy participants to decide whether visually presented tools were effective in various mechanical tasks. Brain activity was measured with 3T BOLD fMRI. We used a multiband EPI sequence with TR = 670 ms. Tools were either highly familiar or unfamiliar according to a preceding behavioral study. On the single subject level we created linear contrasts between the conditions using SPM12.

The group analysis indeed showed higher activation for unfamiliar than for familiar effective tools in the left IFG, left SMG, left MOG and left IOG/ITG. Exploiting our relatively high temporal resolution, we examined the time courses of signicant clusters that had been identified in the whole brain analysis.

Keywords: Tool use; decision; evaluation; left hemisphere

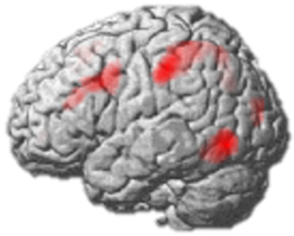


Fig. 1. Overview of all clusters projected on the group brain

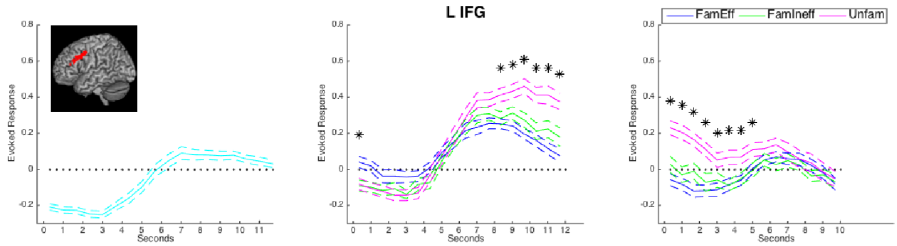


Fig. 2. Time course of the cluster in the left IFG. Cyan curves represent the activation across all conditions during the presentation of the mechanical task.

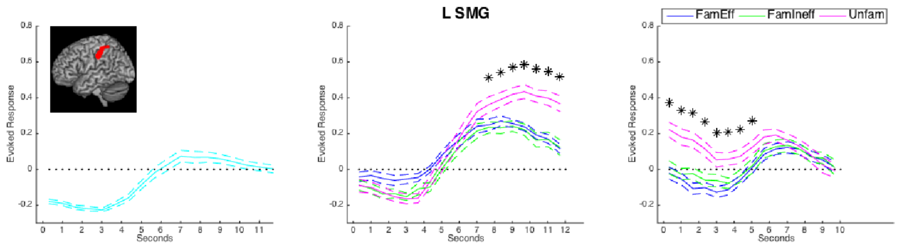


Fig. 3. Time course of the cluster in the left SMG

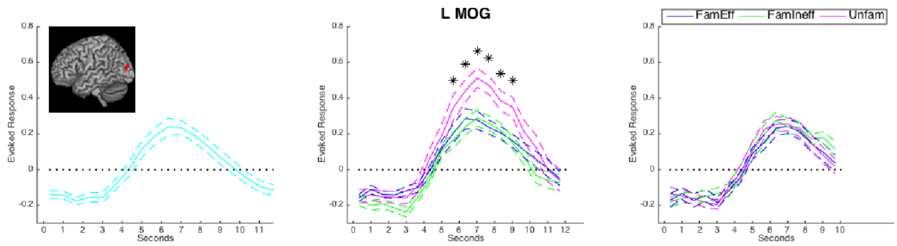


Fig. 4. Time course of the cluster in the left MOG

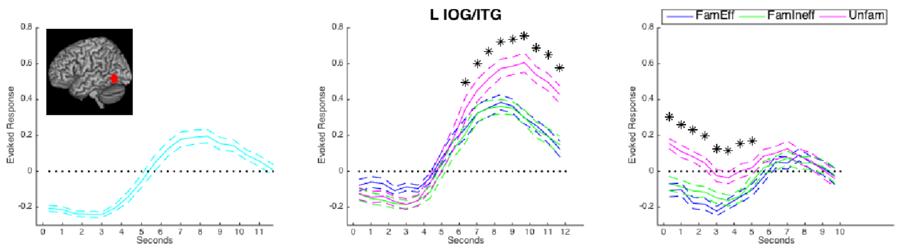


Fig. 5. Time course of the cluster in the left IOG/ITG

Neural Correlates of Semantic Expectation in a Conversation – A Wireless EEG Study of the N400 Effect

Tatiana Goregliad Fjaellingsdal[1], Esther Ruigendijk[2,3], Stefan Scherbaum[4], Martin G. Bleichner[1,2]

1 Department of Psychology, University of Oldenburg, Oldenburg, Germany
2 Cluster of Excellence Hearing4all, University of Oldenburg, Oldenburg, Germany
3 Department of Dutch, University of Oldenburg, Oldenburg, Germany
4 Department of Psychology, Technische Universität Dresden, Dresden, Germany

The understanding of real time language processing (i.e., during conversation) could benefit greatly from increased ecological validity of brain imaging studies [1]. However, conversational setups face several challenges (e.g., control of presentation, timing, and number of turn takes, as well as prevention of artifacts) [2]. As a consequence, to date, conversational implementation during electroencephalographic studies remains minimal [3].

Here we present an approach that allows to study some of the elements of a natural conversation keeping essential constraints for quantitative analysis applying electroencephalography (EEG). The N400 effect, the difference in amplitude between semantically expected and unexpected items, is used here to study language in a conversational setting. The N400 effect has been linked to semantic integration and to prediction building [4]. Therefore, it should not be affected by our role in a conversation (active/passive) since we seem to constantly predict plausible sentence continuations as well as turn-takes in a conversation [5]. Consequently, we expected (1) that semantic violations will lead to an N400 effect and (2) that speaker-switches (as a result of turn-taking during conversation) will not influence the N400 effect, as well as (3) that the N400 effect will not be affected by the role of the person in a conversation: active (that is switch from self-speaker to listener) or passive (that is listening to a speaker who finishes his turn and listening to the new speaker who takes the turn).

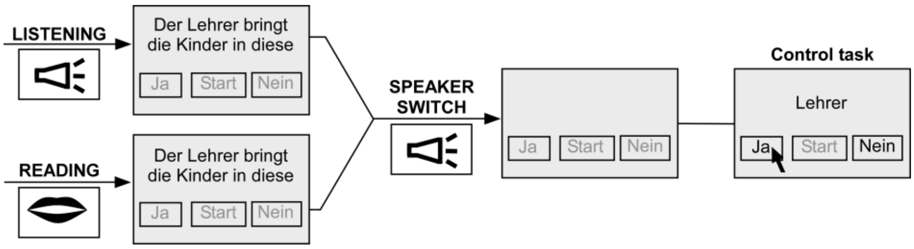


Fig. 1. Screen state of a trial. The first seven words of a sentence are presented. The participant either listens to the pre-recorded sentence fragment read by speaker 'A' (Listening) or reads the sentence fragment out aloud (Reading). The pre-recorded final (eighth) word of the sentence, read by speaker 'B', is then presented (speaker switch, triggered by experimenter). Subsequently, a control task word appears on the screen (task for attention assessing: word present/absent in sentence).
Example sentence (translated): "*The teacher takes the kids into this **school/wound*** (i.e., final word is either congruent/incongruent).

To assess these hypotheses we measured 16 healthy German-native speakers with wireless EEG [6] with an active (reading aloud) and passive (listening) condition (Fig.1) where sentences with semantically congruent and incongruent final words were presented. The interactive element was a speaker-switch for the final word of each presented sentence. Further, correctly categorized words during a control task (word present or absent in sentence) served as attention assessment.

As expected, a significant N400 effect (i.e., larger amplitudes for semantically incongruent than congruent endings) was found with no effect of active (Reading) or passive (Listening) condition on the N400 (Fig.2a). Neither the turn-take nor the conversation role (listener/speaker) seem to affect the semantic integration/expectation. Unexpectedly, a strong modulation of the P200 was also found which was significantly affected by the condition; it was increased for actively reading aloud compared to passive listening (Fig. 2b). This finding can be ascribed to higher order processes of attention and perception which respond strongly to salient, novel stimuli [7]. In the present paradigm this is the speaker switch for every final word. Moreover, this difference is more pronounced for the Reading condition, since there is a switch in sound source (own vs. loudspeaker) and feedback (self vs. externally generated).

Our findings show that the N400 effect is present also during a speaker-switch in a conversation. The speaker-switch, however, seems to influence earlier components of the EEG that are linked to processing of novel auditory stimuli. We conclude that the N400 can be used to study neural correlates of language in conversational approaches.

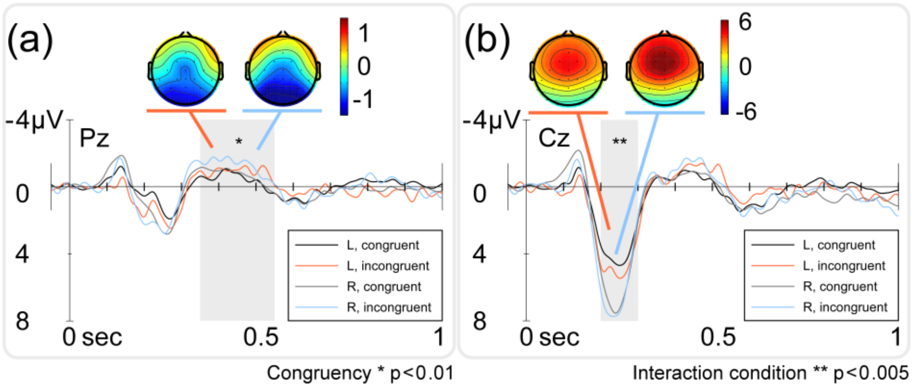


Fig. 2. Grand average ERPs at Pz (a) and Cz (b) with topographies for the incongruent conditions for the Listening (orange) and Reading block (light blue). Zero point is the onset of the final word. Two 2x2 repeated measures ANOVAs with the factors Congruency (congruent/incongruent) and Interaction condition (Listening/Reading) were computed with the mean amplitudes at Pz for the expected N400 effect (350-550ms, gray highlight (a)) and at Cz for the P200 (200-300ms, gray highlight (b)). Significance levels are depicted in the figure. L=Listening, R=Reading.

References

1. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. Behav. Brain Sci. 27, 169–226 (2004).
2. van Berkum, J.J.A.: The Electrophysiology of Discourse and Conversation. In: Spivey, M., McRae, K., and Joanisse, M. (eds.) The Cambridge Handbook of Psycholinguistics. pp. 589–614. Cambridge University Press (2012).
3. Kuhlen, A.K., Allefeld, C., Anders, S., Haynes, J.: Towards a multi-brain perspective on communication in dialogue. In: Willems, R.M. (ed.) Cognitive neuroscience of natural language use. pp. 182–200. Cambridge University Press (2015).
4. Kutas, M., Federmeier, K.D.: Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647 (2011).
5. Levinson, S.C., Torreira, F.: Timing in turn-taking and its implications for processing models of language. Front. Psychol. 6, 1–17 (2015).
6. Debener, S., Minow, F., Emkes, R., Gandras, K., de Vos, M.: How about taking a low-cost, small, and wireless EEG for a walk? Psychophysiology. 49, 1617–1621 (2012).
7. Beauducel, A., Debener, S., Brocke, B., Kayser, J.: On the reliability of augmenting/ reducing: Peak amplitudes and principal component analysis of auditory evoked potentials. J. Psychophysiol. 14, 226–240 (2000).

The interaction dynamics of meta-control parameters and congruency proportion in spatial set shifting

Tobias Grage, Simon Frisch, Stefan Scherbaum

Department of Psychology, Technische Universität Dresden, Dresden, Germany

In goal directed behavior, two opposed constraints have to be satisfied: on the one hand, one has to be able to pursue a goal against obstacles and distractions, thereby avoiding volatility; on the other hand one has to let loose when alternatives become more attractive or the goal becomes unattainable, thereby avoiding futile perseveration. These opposing constraints have been conceptualized as the control dilemma between the shielding of goals from distraction and the shifting of goals when necessary [1]. It is assumed that the balance between perseverating or volatile behavior is regulated by meta-control parameters which configure the cognitive system's default mode of processing [2].

To investigate these meta-control parameters, we introduce a spatial set shifting paradigm in which we shift the balance towards perseverating behavior via congruency proportion [3]. We expect that with such a shift in the balance participants will perform better when it comes to goal shielding and worse when it comes to goal shifting.

In spatial set shifting, two stimuli (i.e. digits) are presented in two of three possible locations (i.e. bottom left, bottom right or top) on a screen. Participants have to categorize (i.e. by number magnitude) the stimulus that they were instructed to attend to (the target) by moving the mouse cursor into a corresponding response box on the top of the screen. After several repetition trials, the positions of target and distracter switch according to specific rules (e.g. the former target position becomes the new distracter position and the previously unused position becomes the new target position).

This task offers two effects to assess both perseverative and volatile behavior: congruency effects (i.e. faster responses for a corresponding target and distracter, indicating the strength of goal shielding) and switch costs (i.e. slower responses after switches of the target position, indicating the strength of goal shifting) [4].

We examine the temporal dynamics of how the different cognitive processes are affected by congruency proportion via mouse tracking [5].

The results corroborate our hypotheses: during blocks with low congruency proportion (20% congruent repetition trials), we found a later onset of the target's influence in switch trials. In repetition trials the effects of low congruency proportion are advantageous in incongruent trials and disadvantageous in congruent trials reflecting conflict adaptation [6].

To put these results in context, we use a dynamic field modeling approach [7]: we match the model's temporal dynamics to those of the experiments and identify neural parameters responsible for the shift towards perseveration.

1. Goschke, T.: Volition in action: Intentions, control dilemmas and the dynamic regulation of intentional control. In: Prinz, W., Beisert, A., and Herwig, A. (eds.) Action Science: Foundations of an Emerging Discipline. pp. 408–434. MIT Press, Cambridge, MA (2013).
2. Gruber, O., Goschke, T.: Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. Acta Psychol. (Amst.). 115, 105–121 (2004).
3. Bugg, J.M., Crump, M.J.C.: In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. Cognition. 3, 367 (2012).
4. Frisch, S., Dshemuchadse, M., Görner, M., Goschke, T., Scherbaum, S.: Unraveling the sub-processes of selective attention: insights from dynamic modeling and continuous behavior. Cogn. Process. 16, 377–388 (2015).
5. Scherbaum, S., Dshemuchadse, M., Fischer, R., Goschke, T.: How decisions evolve: The temporal dynamics of action selection. Cognition. 115, 407–416 (2010).
6. Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D.: Conflict monitoring and cognitive control. Psychol. Rev. 108, 624–652 (2001).
7. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. Psychol. Rev. 109, 545–572 (2002).

Landmark preference during route encoding and retrieval

Karoline Greger[1], Rebecca Albrecht[2], and Rul von Stülpnagel[1]

[1] University of Freiburg
karoline.greger@saturn.uni-freiburg.de
rul.von.stuelpnagel@cognition.uni-freiburg.de
[2] University of Basel
rebecca.albrecht@unibas.ch

Abstract. In this work we extend research on landmark selection by considering the usefulness of landmarks in an actual wayfinding task. To this end, we disentangle the different forms of salience by also considering intersections with only one or none type of salience present. We hypothesized that the interplay between different salience types in one intersection has a positive impact on performance. Results indicate that performance is highest if all types of salience play together or no type of salience is present at all.

Keywords: landmark selection, route recall, landmark salience

People often use landmarks during wayfinding. One conceptualization in research on landmark selection distinguishes two salience types: perceptual and visual salience [1]. A perceptually salient landmark is defined as an object or building with a high visual contrast to its surrounding [2]. A landmark's perceptual salience does not depend on a specific route. In contrast, structural salience is determined by its relative position to a specific route. The combination of both types of salience has been termed joint salience. Previous research on landmark salience focuses on landmark selection without considering actual wayfinding performance. This research shows that people generally prefer the perceptually salient landmark over the structurally salient landmark. In case of a joint salient landmark preferences during selection are even greater [2]. In this work we extend the scope of landmark selection research by focusing on the performance in route recall.

In the experiment we tested 25 participants (13 of them females, 26.41 years, SD = 3.82). Participants had to find a way in a grid-like virtual environment, resembling a urban scenario, consisting of 12 regular intersections, and shown from a first-person perspective. Buildings at the four corners of these intersections served as landmarks. Visual salience was manipulated by all buildings being differently colored (low), or one building contrasting three uniformly colored buildings (high). Structural salience was manipulated by an intersection passed straight (low) or indicating a turn (high). No color was used twice. This approach resulted in five intersection types illustrated in Figure 1. In the first block participants were shown a route through the environment and had to select one building at each intersection considered to be most useful for route recall. In the second block, participants hat to recall the route as correct as possible.

Concerning landmark selection we replicated the findings from [3] for the landmark selection in block 1. The participant's preferences for landmark selection are illustrated in Figure 1. The results indicate a strong preference for the joint salient landmark (Figure 1(A)).

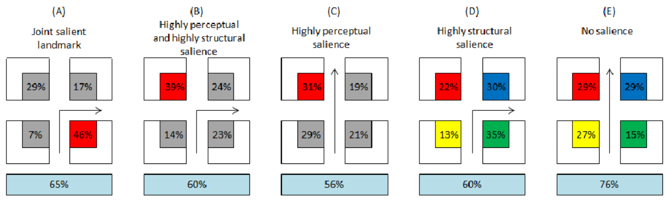


Fig. 1: Landmark selection preferences are presented in percent in the squares' corners. Percentages in the blue boxes below represent correct decisions during route recall.

We compared the success rate of each condition against chance to verify the improvement during wayfinding. In total we found higher error rates. The only intersection type with correct decision rates above chance level were joint salience ($\chi^2(1) = 5.25, p < 0.01$) and when there is no salience ($\chi^2(1) = 15.29, p < 0.01$.

The experiment replicates previous research by emphasizing the exceptional role of a landmark with joint visual and structural salience during landmark selection. This finding extends to route recall in that this type of landmark also helps wayfinding. This results suggest that the more salience types are combined on one object it is not only chosen more likely as a landmark but it also helps route recall. Interestingly, route recall was surprisingly good at intersections with no salient landmarks. We are currently investigating participant's self-reported strategies for explanations of this finding. The preliminary results enable not only insight in landmark selection during an extended condition, but the connection between landmark selection and route recall deserve further analysis.

References

1. Caduff, D., Timpf, S.: On the Assessment of Landmark Salience for Human Navigation. Cognitive Processing 9, 249 – 267 (2008)
2. Röser, F., Krumnack, A., Hamburger, K.: The influence of perceptual and structural salience. Proceedings of the 35th Annual Conference of the Cognitive Science Society (2013)
3. Röser, F.: The cognitive observer-based landmark-preference model: what is the ideal landmark position at an intersection? University of Gießen (2015)

The impact of a humanoid robot's action-selection strategy on humans' perceived naturalness of interaction – A User Study with NAO Playing Rock-Paper-Scissors

Tobias Jakubowitz, André Kowollik, and Ute Schmid

Cognitive Systems Group, University of Bamberg
An der Weberei 5, 96045 Bamberg, Germany
ute.schmid@uni-bamberg.de

For socially interactive robots it is particularly relevant that humans perceive the interaction as natural. Naturalness might be manifested in a humanoid appearance. More relevant are factors like natural behavior and compliance with expected response intervals (Fong, Nourbakhsh, & Dautenhahn, 2003). We are interested in a further aspect which might be relevant for perceived naturalness of human-robot interaction: The impact of the robots cognitive style, that is the strategy by which it selects actions. To our knowledge, this aspect of interaction is researched in human-computer interaction (Young, 2010) but not yet in human-robot interaction.

For our investigation of the influence of cognitive style on perceived naturalness of interaction, we realized different action-selection strategies for the Rock-Paper-Scissors game: (1) **adaptive** – the pattern of the moves of the human from four rounds of the game are used for predicting the next gesture of the human and to select a winning gesture (Cook, Bird, Lünser, Huck, & Heyes, 2011), (2) **random** – one of the moves is chosen randomly, (3) **cheat** – after the human's gesture is recognized, a winning gesture is chosen, (4) **fixed**–a random gesture is selected and kept until the game is over.

In a first study (Kowollik, 2014), we realized the game with the humanoid robot NAO (Fig. 1). Ten subjects played with NAO in individual sessions. A session consisted of four plays where each play consisted of at least three rounds. A round was one when either NAO or the human player had won three moves. The four cognitive strategies were varied within-subjects with one strategy realized

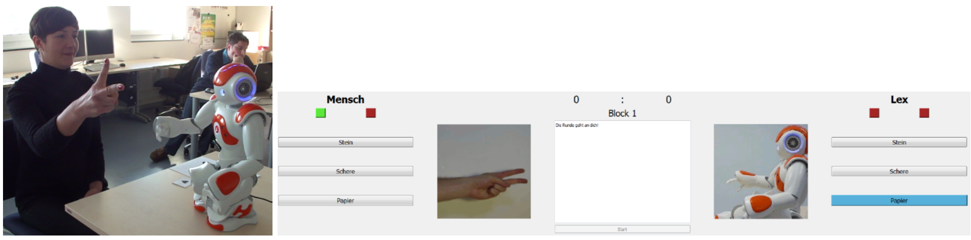


Fig. 1. Interaction with NAO (left) and with computer (right) for Rock-Paper-Scissors

per play in different sequences. After each play, participants answered yes-no-questions about naturalness, human-likeness and fun of the interaction. Results show that naturalness was rated 50% for the random strategy and less for the three others. Human-likeness was rated 70% for the random strategy, 50% for the adaptive strategy, 20% for the cheating, and 10% for the fixed strategy. For both random (80%) and adaptive (70%) participants said that they had fun playing with NAO, for the other two strategies, only 50% of participants said that they had fun. Over all, interactions where NAO realized the random and the adaptive strategy were rated better than the cheating and the fixed strategy.

To have a closer look on the effect of cognitive strategies on perceived naturalness, we conducted an experiment with a computerized version of the game (Jakubowitz, 2015). One version of the game showed a video of NAO as the opponent (see Fig. 1), another version was text-based only. Video vs. text was assigned between-subjects, the four strategies where presented within-subjects as in the previous study. Naturalness and enjoyability were assessed with 7-point Likert scales. Results show no significant differences between the video- and the text-based version of the game, no interaction effect, but a large main effect for strategies for naturalness (Wilks-Lambda: $F = 73.06$; $p \leq .0001$) as well as for enjoyability (Wilks-Lambda: $F = 26.79$; $p \leq .0001$). As in the previous study, the adaptive and the random strategy were rated better than the cheating and the fixed strategy.

Currently, we are preparing a further study with NAO using the same setting as in the computer experiment.

References

Cook, R., Bird, G., Lünser, G., Huck, S., & Heyes, C. (2011). Automatic imitation in a strategic context: players of rock–paper–scissors imitate opponents' gestures. *Proceedings of the Royal Society of London B: Biological Sciences*, *279*(1729), 780786.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, *42*(3), 143–166.

Jakubowitz, T. (2015). *The impact of a humanoid robot's action-selection strategy on humans' perceived naturalness of interaction – A user study with nao playing rock-paper-scissors.* Bachelor thesis in psychology, University Bamberg.

Kowollik, A. (2014). *Mensch-Roboter-Interaktion in gestenbasierten Zweipersonenspielen: NAO spielt Schere, Stein, Papier.* Master thesis in computer science, Universities Bamberg and Erlangen.

Young, S. (2010). Cognitive user interfaces. *Signal Processing Magazine*, *27*(3), 128–140.

Semantics of Persian Spatial Term ǰelo Based on Principled Polysemy Model

Marjan Daneshvar Kashkooli

Isfahan University

1 Abstract

This study investigates the semantic aspects of the polysemous Persian spatial term, ǰelo, within a cognitive linguistics framework. We seek to identify how Persian speakers conceptualize space in terms of contexts where ǰelo ('in front of') is applied. To this end, the 'Principled Polysemy Model' developed by Tyler and Evans (2003) is employed as the main analytical tool to construct its semantic network. Our analysis is based on the data taken from various written and spoken texts from contemporary Persian by concentrating on different applications of this specific term. One key purpose of this research is to find out how polysemy can be accounted for as a non-arbitrary motivated phenomenon. In other words, we attempt to investigate the polysemy apparatus of ǰelo in order to show how cognitive motivations trigger its polysemous configuration. Moreover, the study is geared towards determining whether the Principled Polysemy Model originally developed for English prepositions is compatible to a distant language, namely Persian. Semantic analysis of ǰelo finally depicts how polysemy is driven by certain cognitive motivations. Additionally the study reveals that as long as ǰelo is used as a single morpho-syntactic word, the model is capable of accounting for its semantics in terms of a polysemy network consisting of a primary sense plus four distinct senses. However, in the phrasal constructions such as ǰelo-giri kardan ('to prevent') that encode abstract concepts, the model does not supply a conclusively efficient tool. Hence, we argue that regarding abstract uses in such multi-word constructions, the suggested systematic principles cannot practically help determine the distinct senses extended from the primary sense, and the challenge of sense distinction remains unresolved.

2 Theoretical Framework and Method

In order to identify the semantic network associated with Persian ǰelo, this study follows procedures suggested in Tyler & Evans' Principled Polysemy Model. This consists of two stages of primary sense representation as Proto-scene, and the semantic network of other extended senses.

2.1 Proto-scene (the primary sense)

According to this model, a major step towards the precise analysis of polysemous expressions is to identify the central (primary) sense schematically represented as

Proto-scene. Proto-scene serves as an image schema built by abstracting across the scenes associated with one spatial preposition. The model provides several steps to arrive at the proto-scene, among which the following steps are taken in the present study.

1. Attending to the etymological roots of the word and looking for the earliest attested meaning.
2. Considering the spatial configuration of F & G in variety of the term's occurrences and looking for one that is least complicated and most applicable to various uses.
3. Paying attention to the contrastive sets the term participates in, which can lead us to a better candidate for the term's primary meaning.
4. Taking into consideration the morphological presence of the term in derived or more complex words to see what sense of the term shows higher degree of flexibility and frequency.

2.2 Polysemy network

After the proto-scene is devised, the next step is identification of distinct but related senses extended from the primary sense that will eventually provide the polysemy network. Principled Polysemy Model proposes two criteria to ensure a different use corresponds to a distinct sense. The first is that the spatial term must have non-spatial meaning or one spatial sense different from the proto-scene. The second is that context-independent uses should occur, i.e. uses where meaning is not computable from knowledge of another sense or the context. An example of this would be English over in 'There were over a hundred people waiting outside!'. Here the two criteria are met for judging the 'more' sense of over as a distinct meaning extended from its primary spatial sense. 'More' sense is non-spatial on one hand, and cannot be directly computed from other words around it on the other.

As far as the data are concerned, attested Persian sentences from written or spoken language will be used in the present study. The instances are selected from newspapers, books, web pages, and dictionaries, and from the common daily utterances produced by Persian native speakers.

3 Conclusion

This study, with a cognitive approach to semantics, set out to discuss how one Persian spatial term is treated as a complex category with multiple meanings. This analysis followed the cognitive linguistics' claim that meanings are organized as polysemy networks in which senses are systematically (not arbitrarily) extended from one primary sense. Such analysis draws on cognitively motivated notions and processes including construal, conceptualization, and experiential correlations that all are suggested for motivating and explaining the meaning extension involved. (Lakoff 1987, langacker 1987, 1991) Analysis of data that composed of occurrences of ǰelo in various written or spoken sources showed that spatial term ǰelo stands for a complex category when it comes to its semantic aspects. Additionally, the study highlighted experiential grounds as motivation behind ploysemous nature of the term under focus. This, in turn, signaled the embodied nature of semantic extension. Application of the Principled Polysemy Model at the first step resulted in the primary sense of the term schematically represented in the form of proto-scene. On the second step, the polysemy network was sketched with four distinct senses derived from the proto-scene, which generally covered the term in its occurrences as an independent morpho-syntactic word. Although both these steps confirmed that, the followed procedures taken from Tyler & Evans' model could be applied to ǰelo, the analysis of many abstract uses in some longer than one word constructions was not flawless for distinct sense discrimination. Following the model's two criteria could not help making sense-discrimination judgments in a coherent way. Not only this challenge remained unresolved, we started to wonder if the theoretically attractive second criterion in itself, i.e. context-independency for a distinct sense, is rationally sensible and practically efficient in the first place. In general, the findings of this study can be interpreted briefly as follows: the model seems to provide a partially replicable framework that caters for primary sense and polysemy network at the level of single-word applications of Persian spatial term ǰelo. However, for some multi-word verbal constructions where the term in question occurs, the two criteria suffer uncertainty and lack of coherency that implies the need for the researcher's more or less subjective decision. All in all, the study suggests that on the level of constructional aspects of its semantic analysis the model calls for some revision and modification, beside acknowledging the fact that more investigations on variety of linguistic data would be necessary for its better and more accurate evaluation.

Challenging the distinction between presupposition holes and plugs

BACKGROUND: Sentences like (1a) are generally assumed to have a presupposition, given in (1b), which is due to the factive predicate 'be glad'. Such predicates are labeled *presupposition triggers* (see Beaver & Geurts 2013 for a recent overview).

(1) a. Jonas is glad that Bill is in town.
    b. >> Bill is in town.

Since Karttunen (1973), it is assumed that presuppositions 'project' past certain operators (so-called "holes"), which include modal adverbs (e.g. *perhaps*): a presupposition that is triggered in the scope of *perhaps* still has to be met in the 'global' utterance context. To illustrate, (2a) shares the presupposition of (1a), as shown in (2b).

(2) a. **Perhaps** [Jonas is glad that Bill is in town].
    b. >> Bill is in town.

By contrast, there are operators ("plugs") that stop presuppositions from projecting. Such plugs include verbs of saying. To illustrate, the common assumption is that (3a) no longer triggers a global presupposition that Bill is in town (as indicated by (3b)).

(3) a. Markus **told** Hanna that [Jonas is glad that Bill is in town].
    b. >/> Bill is in town.

The distinction between "holes" and "plugs", while widely accepted, has received limited attention in the experimental literature, giving rise to the question of whether naïve speakers behave as expected (in (2b)/(3b)) when encountering such utterances.

METHOD: In an offline ratings experiment (40 items in 2x2 design), we tested whether native speakers of German are sensitive to the "hole"/"plug" distinction. A sample item is given in (4), structured as follows. Participants (n=45) read a German sentence that contained one of five presupposition triggers (factives, iteratives, aspectual elements, temporal clauses, *it*-clefts), as well as a hole vs. a plug. Subsequent questions asked the following: how certain is it (on a 7-point scale) that the presupposed proposition is true? We controlled for projection behavior by manipulating the type of question. In the 'global' condition, (4a-b), the question is unmodified, whereas in the 'local' condition, the question is modified in order to ask about whether the presupposition holds *within* the scope of the relevant ('plug'/'hole') operator, as shown in (4c-d).

(4) a. **[plug + global]** *Markus erzählte Hanna, dass Jonas sich darüber freue, dass übermorgen die Büchermesse beginnt.*
    'Markus told Hanna that Jonas is glad that the book fair will start the day after tomorrow.'
    Wie gewiss ist es, dass übermorgen die Büchermesse beginnt?
    'How certain is it that the book fair will start the day after tomorrow?'

    b. **[hole + global]** *Vermutlich freut sich Jonas darüber, dass Übermorgen die Büchermesse beginnt.*
    'Probably Jonas is glad that the book fair will start the day after tomorrow.'
    Wie gewiss ist es, dass Übermorgen die Büchermesse beginnt?
    'How certain is it that the book fair will start the day after tomorrow?'

c. **[plug + local]** *Markus erzählte Hanna, dass Jonas sich darüber freue, dass übermorgen die Büchermesse beginnt.*
‘Markus told Hanna that Jonas is glad that the book fair will start the day after tomorrow.’
Angenommen, Markus sagte die Wahrheit: wie gewiss ist es dann, dass übermorgen die Büchermesse beginnt?
‘Assume that Markus spoke the truth: If so, how certain is it that the book fair will start the day after tomorrow?’

d. **[hole + local]** *Vermutlich freut sich Jonas darüber, dass Übermorgen die Büchermesse beginnt.*
‘Probably Jonas is glad that the book fair will start the day after tomorrow.’
Angenommen, diese Vermutung erweise sich als richtig: wie gewiss ist es dann, dass übermorgen die Büchermesse beginnt?
‘Assume that this conjecture turns out to be true: If so, how certain is it that the book fair will start the day after tomorrow?’

PREDICTIONS: The traditional view (Karttunen 1973) predicts that the [hole+global] condition is rated higher than the [plug+global] condition (while [hole+local] and [plug+local] may receive similar ratings) (= **Prediction 1**). Moreover, since a reader does not know whether, say, Markus spoke the truth (in (4a/c)) or whether the conjecture that *vermutlich* ‘probably’ marks is true (in (4b/d), we expect that ‘local’ questions give rise to higher certainty ratings (as participants judge the presupposition as it would be without such modification) than the ‘global’ questions (= **Prediction 2**).

RESULTS: Our experiment yields a main effect of question type ($p<0.001$), with higher ratings for the local conditions than for the global conditions, as shown in *Fig. 1* and *Fig. 2*.

Figure 1	**global**	**local**
hole	5.38	5.78
plug	5.42	5.72

There is no interaction ($p=0.52$) and no main effect of the hole vs. plug distinction ($p=0.88$).

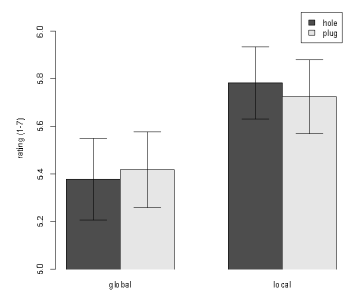


Figure 2

DISCUSSION: Our results are compatible with **Prediction 2**, indicating that subjects were sensitive to the global/local distinction and that they carried out the task correctly. However, our results are incompatible with **Prediction 1**, thus challenging the idea that there is a hole/plug asymmetry that is accessible to naïve native speakers. Our findings are relevant for future studies on the processing of presupposition projection, since the hole/plug asymmetry is typically tacitly assumed to be a firmly established fact.

REFERENCES:
Beaver, D. & B. Geurts. 2013. Presupposition, *Stanford Encyclopedia of Philosophy* (Fall 2013 Ed.), <http://plato.stanford.edu/archives/fall2013/entries/presupposition/>.
Karttunen, L. 1973. Presuppositions and compound sentences. *LI* 4, 169-193.

Learning to cope with uncertainty during the acquisition of mental models

Johanna Renker, Gerhard Rinkenauer

Leibniz Research Centre for Working Environment and Human Factors
renker@ifado.de

Keywords: task uncertainty, mental representation, learning model

1 Introduction

The acquisition of a mental model can be considered as a learning process, viz. information is successively stored in memory during the performance of a cognitive task (Speelman, Kirsner 2008). The learning process as well as the mental model development can be improved by practicing and is often part of category-based induction (Chen et al. 2016). However, the mental model is a reduced representation of an uncertain reality. Strategies to cope with the existing uncertainty are for example searching for relevant information and ignoring irrelevant information (Lipshitz, Strauss 1997). In a recent study Renker and Rinkenauer (2016) employed eye movement patterns to assess the information gathering during the acquisition of a mental representation under uncertainty. The findings revealed intensive visual search behavior at the beginning which decreased considerably during the acquisition process even the objective task uncertainty did not change. The current work aims to extend the findings of the recent study and focuses directly on the learning process of an uncertain task. Two questions are addressed in the following: How fast do participants develop a mental representation under uncertainty? How do participants represent the objective task uncertainty?

2 Method

To study the development of mental models under uncertainty a new experimental task was developed: the occluded visual spatial search task (OVSST). This task consists of a prediction task and a reaction task. In this study we will only focus on the prediction task. Participants were instructed to predict at which of three exits (left, top, right) one of three objects (circle, triangle or square) will reappear out of a dark room. In order to improve task performance participants had to learn the underlying probability concept of the OVSST. Every object was associated to one of the exits with a higher probability (74%) and to the other two exits with a lower probability (11%). Participants had no prior knowledge about this probability structure. In the remaining cases (4%) a rare occurrence happened that was unpredictable for the participants viz. the objects reappeared at the bottom entrance. Every participant performed 324 trials of the OVSST subdivided into four blocks of 81 trials each.

3 Results

Results of repeated measures analyses of variance showed an increasing number of correct predictions over blocks (Block1: *M*=47.11, *SD*=9.72; Block4: *M*=54.78, *SD*=8.75). However, only the increase from Block1 (*M*=47.11, *SD*=9.72) to Block2 (*M*=53.44, *SD*=8.75) was significant $F(1;17)=17.14$, $p=.001$. Averaged number of correct and incorrect predictions were fitted to two learning models: A power model and an exponential model. The power model revealed a better fit (R^2=.92) than the exponential model (R^2=.89) presumably due to the analysis of several individuals (Speelman, Kirsner 2008). The prediction data revealed that the main tendencies of the probability structure were learned within the first 20 trials. Participants predicted already at this early learning state in 63.7% of the cases the likely exits and in 18.2 % of the cases the unlikely exits. In the last block participants chose with 92.5% the likely exits and with 3.8% the unlikely exits. Furthermore, the asymptote of the power function approaches 99.26% for the likely exits and 0.45% for the unlikely exits. Thus, the model fit suggests that participants seem to ignore the objective task uncertainty.

4 Discussion and Conclusion

Learning a probability concept as provided by the OVSST seems to need only few repetitions presumably because relations are easy to memorize. However, participants seem to ignore lower probabilities and develop a response strategy that considers only the association between the object and the likely exit, viz. in their mental model uncertainty seems to be redefined somehow as certainty. This strategy is consistent with the ignorance if irrelevant information as a coping strategy mentioned earlier (Lipshitz, Strauss, 1997). In fact, participants obviously do not represent the objective uncertainty of the OVSST contrary to the concept of probability matching presumably due to the reduced task complexity. However, the results are consistent with findings of Edwards (1961) indicating more extreme predictions than actual occurrences.

References

1. Chen, Stephanie; Ross, Brian; Murphy, Gregory (2016): Eyetracking Reveals Multiple-Category Use in Induction. In *Journal of Experimental Psychology: Learning, Memory and, Cognition,* pp. 1-18. DOI: 10.1037/xlm0000222.
2. Edwards, Ward (1961): Probability Learning in 1000 Trials. In *Journal of Experimental Psychology* 62 (4), pp. 385-394. DOI: 10.1037/h0041970.
3. Lipshitz, Raanan; Strauss, Orna (1997): Coping with Uncertainty. A Naturalistic Decision-Making Analysis. In *Organizational Behavior and Human Decision Processes* 69 (2), pp. 149–163. DOI: 10.1006/obhd.1997.2679.
4. Renker, Johanna; Rinkenauer, Gerhard (2016): The Acquisition of Mental Representations under Uncertainty: An Eye Movement Study (submitted).
5. Speelman, Craig; Kirsner, Kim (2008): Beyond the Learning Curve. The construction of mind. Oxford: Oxford Univ. Press (Oxford Scholarship Online).

Dissociating components of cognitive flexibility in semantic space: Continuous measures, dynamic modeling and clinical assessment

Stefan Scherbaum, Tobias Grage, Moritz Walser, Katrin Hummel,
Maja Dshemuchadse

Cognitive flexibility as an executive sub function is often studied as the ability to shift cognitive sets [1]. We propose to further distinguish two components of cogniti-ve flexibility: shifting flexibility and spreading flexibility. These components base on independent underlying mechanisms: shifting flexibility refers to a complete shift between unrelated sets, while spreading flexibility refers to the breadth of spreading activation between related sets.

To separate these two components, we used a homonym relatedness judgement task [2]. In this task, participants have to judge whether a target word (e.g. drummer) is related to a second word (the so-called associate, e.g. stick). In the critical so-called homonym trials, the target word is a homonym (e.g. band) and the associate can either be related to the meaning of the homonym (e.g. a group of musicians) that has been primed by a previous trial or to the meaning of the homonym (e.g. ribbon/strip) that has not been primed. Furthermore, the associate can be strongly related to the homonym (e.g. singer) or only loosely connected (e.g. pitch). Hence, the task tested shifting flexibility (primed vs. unprimed meaning of the homonym) and spreading flexibility (strongly or weakly related associate). We combined this task with mouse tracking [3] and found in a study with 20 participants that these two components of cognitive flexibility indeed follow independent temporal patterns during relatedness judgments.

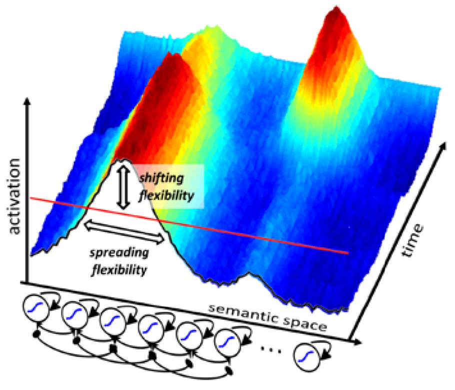


Fig. 1. Dynamic neural field based model with shifting flexibility reflected in the stability of peaks of neural activation and spreading flexibility reflected in the width of neural activation peaks.

adfa, p. 1, 2011.

These results are in concordance with the predictions of a dynamic neural field based model [4] that, first, assumes the independence of the two components of flexibility and, second, proposes two neural parameters to implement the two components (See Fig. 1). The model represents associative distance in a continuous semantic space within a continuous neural field. Shifting flexibility is implemented as the stability of peaks in such a field and hence supports long-range jumps in this semantic space when stability at a location in this field is low. Spreading flexibility is implemented by the width of peaks in the field and hence supports short-range moves within a certain area of the field when peak width is high [5].

To validate the model, we derived parameter-specific predictions related to symptoms of obsessive compulsive disorder as it has been done previously with related models [6]. We expected that OCD primarily loads on the first component's parameter, that is decreased shifting flexibility in OCD leading to perseveration within an associative context [7]. From a screening sample of 808 students, we selected 54 students that exhibited either low scores or high scores in the Obsessive-Compulsive Inventory [OCI-R - 8]. This sample of subclinical students performed the homonym relatedness judgement task, yielding the expected pattern of impaired shifting flexibility in the high OCI group.

Taking together, we propose that studies about cognitive flexibility in the area of executive functions should take the two independent components into account, especially when studying moderators of cognitive flexibility.

1. Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., Wager, T.D.: The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. Cognit. Psychol. 41, 49–100 (2000).
2. Dshemuchadse, M.: Einfluss von Stimmungen auf kognitive Parameter. Südwestdeutscher Verlag (2009).
3. Scherbaum, S., Dshemuchadse, M., Fischer, R., Goschke, T.: How decisions evolve: The temporal dynamics of action selection. Cognition. 115, 407–416 (2010).
4. Erlhagen, W., Schöner, G.: Dynamic field theory of movement preparation. Psychol. Rev. 109, 545–572 (2002).
5. Dshemuchadse, M., Grage, T., Scherbaum, S.: Action dynamics reveal two components of cognitive flexibility in a homonym relatedness judgement task. Front. Cogn. 6, 1244 (2015).
6. Rolls, E.T., Loh, M., Deco, G.: An attractor hypothesis of obsessive-compulsive disorder. Eur. J. Neurosci. 28, 782–793 (2008).
7. Jelinek, L., Hauschildt, M., Hottenrott, B., Kellner, M., Moritz, S.: Further evidence for biased semantic networks in obsessive–compulsive disorder (OCD)J. Behav. Ther. Exp. Psychiatry. 45, 427–434 (2014).
8. Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., Salkovskis, P.M.: The Obsessive-Compulsive Inventory: Development and validation of a short version. Psychol. Assess. 14, 485–496 (2002).

LANGUAGE CUES IN THE FORMATION OF HIERARCHICAL REPRESENTATIONS OF SPACE

Wiebke Schick[1,*], Marc Halfmann [2], Gregor Hardiess and Hanspeter A. Mallot [2]
[1], [2] Cognitive Neuroscience, Dept. of Biology, University of Tübingen, Auf der Morgenstelle 28, Tübingen
*Corresponding author: wiebke.schick@student.uni-tuebingen.de

Background:
The formation of hierarchical representations of space can be induced by the spatial adjacency of places marked with landmark objects from the same semantic category, as was demonstrated in a route planning experiment (Wiener & Mallot, 2003).
Aims:
Using the same paradigm, we tested the impact of linguistic cues with various hierarchical categorization principles on inducing region perception. In five conditions, places of the experimental environment were characterized (i) with landmark objects (*Landmark)*, (ii) with city names and inner-city locations (*City names*), (iii) with city names derived from three places of the former condition (*Pars-pro-toto*), (iv) with nouns from different semantic categories (*Semantic*) and (v) with locations from multicomponent institutions (*Compounds*).
Method:
In a virtual environment, subjects performed navigation tasks (*Travelling -Salesman -Problem*). All mazes consisted of 12 places, six of them arranged in a hexagon (*iterated y-maze,* Fig.1). This structure allows for equidistant, but region-sensitive alternatives of the test routes: one alternative transgresses more regions – and therefore regional boundaries – than the other one. There are no visible, physical boundaries. *We expect a region effect, namely a* preference for the routes passing fewer regions, if the categorisation schemes of the cues induce a perception of regional subdivsion.

Figure 1: *Left:* Survey of the experimental environment. The regions are depicted through different shades of grey. *Center:* Experimental setup, *Landmark* condition. *Right:* Cue representation (*Semantic* condition).

Results:
The results of the *Landmark* condition confirmed the findings by Wiener & Mallot. For the linguistic conditions, higher error rates as well as strong differences in the prevalence of region-consistent route choices were found. A significant preference was found only for the *Compounds* condition (Fig.2).
As a measurement for the associations between the linguistic cues, a Latent Semantic Analyis was performed with the texttiling-algorithm. The measure is called PmiLR, uses Pointwise Mutual Information, and yields information statistics for the degree of synonymity and the likelihood of co-occurrence of a given word pair (Turney, 2001).

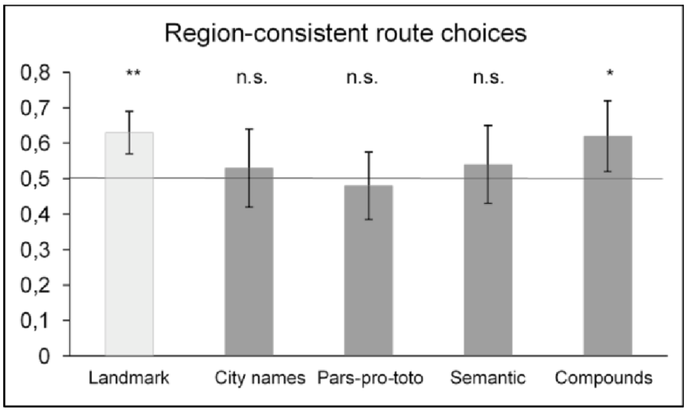


Figure 2: *Route-choice preferences: Proportion of the correct routes that were region-consistent (including chance level and standard deviation).*

A visualization through multidimensional scaling of the computed distances showed a correspondence to the intended regional categorization scheme only for the *Compounds* condition (Fig. 3).

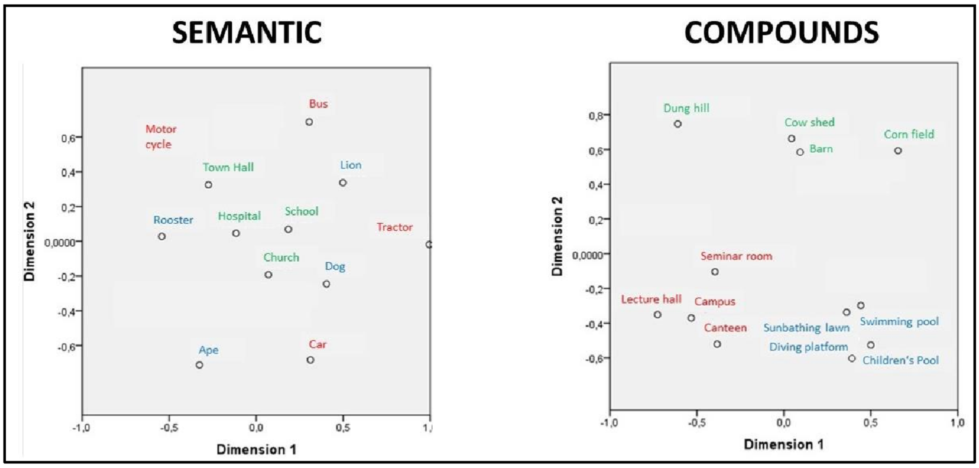


Figure 3

Display of the contextual associations between the linguistic cues computed through the PmiLR-analysis of the place names. Left: Semantic; right: Compounds.

Conclusions:

(i) While objects of different semantic categories were used to establish a regionalized spatial code, the names of such objects alone did not induce a region effect. This indicates that language-based regionalization requires additional cues.

(ii) A privileged status that influences and affects cognition cannot be assigned nominally.

(iii) Linguistic landmarks can be constructed, but have to contain additional associations. For establishing a spatial representation as well as for memorizing place names, consideration of context and of well-encoded, stereotypical scenarios can facilitate region formation and place recall.

References

- Turney, P. (2001). Mining the web for synonyms: PMI-IR versus LSA on TOEFL. NRC Publications Archive.
- 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.

Co-representation of Others' Spatial Task Constraints in Joint Action

Laura Schmitz, Cordula Vesper, Natalie Sebanz, and Günther Knoblich

Department of Cognitive Science, Central European University, Budapest, Hungary
Schmitz_Laura@phd.ceu.edu, VesperC@ceu.edu, SebanzN@ceu.edu, KnoblichG@ceu.edu

Keywords: joint action · co-representation · coordination · social cognition.

1 Abstract

In the present study, we asked whether individuals engaged in a joint action rely on task co-representation to achieve coordination, even when this implies increased movement effort. The results of four experiments showed that unconstrained actors represented a co-actor's spatial task constraint, adjusting their own movements accordingly. These findings suggest that joint action partners rely on task co-representation for coordination.

2 Present Study

Previous research has shown that when acting independently alongside others, individuals often represent each other's tasks [1]. In joint actions, task co-representation can be effective in facilitating joint action coordination [2], but individuals can also jointly achieve a goal by using coordination processes that do not require them to represent another's task [3, 4]. The goal of the present study was to examine whether joint action partners engage in task co-representation to achieve interpersonal coordination even when there are alternative coordination processes that imply less movement effort. We devised a joint movement task [5] where co-actors could either represent their partner's task constraint at the cost of performing more effortful movements or simply slow down their own actions to match the predicted duration of the partner's actions.

Pairs of participants performed reaching movements back and forth between two targets, synchronizing their landing times. The movements of one actor were sometimes constrained by an obstacle obstructing her movement path while the other actor's movements were unconstrained. We predicted that the unconstrained actor's movements should be higher in trials where the co-actor's movement is constrained by an obstacle than in trials where the co-actor's movement is not constrained, indicating an effect of representing the co-actor's spatial constraint on the unconstrained actor's movements.

The results of four experiments reliably showed that unconstrained actors represented their co-actor's task constraint such that they increased their own movement amplitude when their co-actor moved over an obstacle. These effects were considerably larger when coordination demands were higher (Experiment 2) and occurred irrespective of visuospatial perspective (Experiment 3). Experiment 4 suggested that unconstrained actors represented the object property constraining the co-actor's movements rather than specific parameters of these movements. We conclude that task co-representation for coordination seems to be a preferred way of achieving interpersonal coordination even if it implies increased movement effort.

Finally, it is an open question whether task co-representation occurred spontaneously in the present study, as suggested by previous findings [1], or whether participants *intentionally* formed and maintained a representation of the co-actor's task constraint and *intentionally* modulated their own movements accordingly.

3 References

1. Sebanz, N., Knoblich, G., Prinz, W.: Representing others' actions: just like one's own?. Cognition, 88(3), B11-B21 (2003)
2. Vesper, C., van der Wel, R. P., Knoblich, G., Sebanz, N.: Are you ready to jump? Predictive mechanisms in interpersonal coordination. Journal of Experimental Psychology: Human Perception and Performance, 39(1), 48-61 (2013)
3. Vesper, C., Butterfill, S., Knoblich, G., Sebanz, N.: A minimal architecture for joint action. Neural Networks, 23(8), 998-1003 (2010)
4. Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., Schmidt, R. C.: Rocking together: Dynamics of intentional and unintentional interpersonal coordination. Human Movement Science, 26(6), 867-891 (2007)
5. van der Wel, R. P., Fu, E.: Entrainment and task co-representation effects for discrete and continuous action sequences. Psychonomic bulletin & review, 22(6), 1685-1691 (2015)

I-CARE: Individual Activation of People with Dementia

Tanja Schultz[1], Felix Putze[1], Timo Schulze[1], Ralf Mikut[2], Wolfgang Doneit[2], Andreas Kruse[3], Anamaria Depner[3], Ingo Franz[4], Marc Aurel Engels[5], Philipp Gaerte[5], Dietmar Bothe[6], Christof Ziegler[6], Irene Maucher[7], Michael Ricken[8], Todor Dimitrov[8], Joachim Herzig[8], Keni Bernardin[9], Tobias Gehrig[9], Jana Lohse[10], Marion Adam[10], Monika Fischer[10], Massimo Volpe[10], and Clarissa Simon[10]

[1]Cognitive Systems Lab, Universität Bremen
tanja.schultz@uni-bremen.de

[2]Institut für Angewandte Informatik, Karlsruher Institut für Technologie
[3]Institut für Gerontologie, Universität Heidelberg, [4]Diakonische Hausgemeinschaften Heidelberg e.V., [5]Media4Care GmbH, Berlin, [6]topsystem Systemhaus GmbH, Würselen, [7]Deutsche Telekom Healthcare and Security Solutions GmbH, [8]Anasoft Technology AG, Bochum, [9]Videmo Intelligente Videoanalyse GmbH & Co. KG, Karlsruhe, [10]AWO Karlsruhe gemeinnützige GmbH, Karlsruhe

1 Motivation

Many developed countries see a dramatic demographic change. Their population is aging, the life expectancy increases while birthrate decreases. Since the risk of cognitive decline and dementia is drastically increasing with age, Europe expects that the amount of people developing dementia over the next 20 years will double every 5 years. Consequently, the demand of care is substantially growing and expected to be provided by professional as well as informal carers.

We are convinced that technical systems are very suitable to tackle some of the oncoming challenges. Unfortunately, the number of technical systems available for care of people with dementia is still rather limited [1]. Intelligent technical systems that automatically adapt to biographic information, details of past sessions and online recorded sensor signals may allow highly individualized therapy concepts. Such systems could relieve professional as well as informal carers.

The recently started project I-CARE* with seven interdisciplinary partners in academia and industry as well as social services aims at the development of technical innovations in human-computer interaction to support the care of people with dementia in our aging society. In particular, I-CARE focuses on technical support to ease the burden on relatives and professional carers. This is envisioned to be achieved by an adaptive and mobile technical system, which activates and promotes individual cognitive, social, and motor skills.

---

* I-CARE project homepage, https://www.projekt-i-care.de

2 The I-CARE System

We envision the adaptive and mobile I-CARE system, that learns about the individual needs and potentials of people with dementia and facilitates the building of ad-hoc activation groups. The I-CARE system will provide individualized activation content to informal caregivers by analyzing individual activation needs, potentials and daily condition of people with dementia.

I-CARE will be equipped with a *recommender* system, which preselects and suggests a small subset of appropriate items. These suggestions are learned based on explicit or implicit preferences, i.e. individual biographic information, graphical and voice-based ratings made by the user, stress, emotions, or other behavioral user reactions. At deployment, the recommender system has no interaction history for any user, thus preferences are unknown. For proper initialization, information are taken from biographic data, thus semantic similarity [3] and item similarity will be employed.

The system runs on a tablet computer with a straight-forward, easy to use, and intuitive user interface. The tablet camera will be applied for face detection and the identification of emotions from facial expressions, while the microphone allows for voice activity and emotion analysis. In addition to the tablet, users will be equipped with an unobtrusive wrist band. The device measures motion based on inertial sensors, electrodermal activity (EDA) and the cardiac signal, i.e. heart rate (HR) and heart rate variability (HRV). While the first will be applied to interpret the user's physical activities, the latter two will serve to differentiate stressful from relaxing situations [4] during an activation session. By combining the information from the tablet and wrist device, we hope to obtain detailed information on the affective state and engagement of users. Among others, this information will also be applied to the recommender system as implicit rating.

Acknowledgments: The project I-CARE "Individuelle Aktivierung von Menschen mit Demenz" is funded by the Federal Ministry of Education and Research (BMBF) within the research programme "IKT 2020 - Forschung für Innovationen" as joint project under the reference number V4PID062.

References

1. Schultz, T., Putze, F., Mikut, R., Weinberger, N., Boch, K., Schmitt, E., Decker, M., Lind-Matthäus, D., Metz, B.R.: Technische Unterstützung für Menschen mit Demenz - Ein Überblick. In: Schultz, T., Putze, F., Kruse, A. (eds.) Technische Unterstützung für Menschen mit Demenz, pp. 1–18. KIT Scientific Publishing (2014)
2. Kruse, A. (ed.): Lebensqualität bei Demenz? Zum gesellschaftlichen und individuellen Umgang mit einer Grenzsituation im Alter. Akademische Verlagsgesellschaft, Heidelberg (2010)
3. Liu, H., Singh, P.: ConceptNet - a practical commonsense reasoning tool-kit. BT Technology Journal 22(4), 211–226 (2004)
4. Picard, Rosalind W.: Affective computing. MIT press, Cambridge (1997)

Let's decide together! Joint delay decision-making improves delay discounting

Diana Schwenke[1], Maja Dhesmuchade[1], Cordula Vesper[2] , Martin G. Bleichner[3], Stefan Scherbaum[1]

1 Department of Psychology, Technische Universität Dresden, Dresden, Germany
2 Department of Cognitive Science, Central European University, Budapest, Hungary
3 Department of Psychology, University of Oldenburg, Oldenburg, Germany

Important decisions often relate to choices between immediate or delayed gains, e.g. spending money for short-term enjoyment or saving for long-term investment. The phenomenon that people tend to devaluate later gains is referred to as delay discounting [1]**.** In everyday life, such choices often depend on the joint agreement of multiple people. Research on group decision-making demonstrated a group advantage over the individual performance of its members on a variety of decision making tasks [2,3]**.** Despite this remarkable amount of insights, the question of how two people jointly evaluate delay discounting situations has not been addressed empirically. Therefore, we investigate, first, whether delay discounting benefits from dyadic decision-making and second, how their choices may become more effective.

To assess these questions we developed a novel task in which participants executed a sequence of choices between a sooner but smaller (SS) or a later but larger (LL) delivered reward in an individual and in a dyadic decision making condition, executing choices by navigating a curser via joystick movement. While in the individual condition, participants moved their cursor directly to the target box associated with the chosen option; in contrast, they had to coordinate their movements and therefore negotiate their preferences with their partner in the dyadic condition. With this, we were able to track the individual decision in the individual condition, the initial individual decision within the dyadic condition and the final dyadic consent.

Pairwise comparison of these three levels of decision-making revealed that the final dyadic decision resulted significantly less often in a sooner-but-smaller (SS) choice compared to the individual decision and the pre-decision, revealing a relative reduction of delay discounting. Furthermore, when analyzing the quality of choices based on a normative choice model [4], we found clear evidence that dyads outperformed individuals concerning individual decision-making and the initial decision of each participant in the dyadic condition.

In conclusion, our findings support the idea that delay discounting benefits from social collaboration and furthermore identified dyadic interchange [5] rather than social facilitation [6] as a mechanism to improve delay discounting decision.

References:

[1] G. S. Berns, D. Laibson, and G. Loewenstein, "Intertemporal choice-toward an integrative framework," *Trends Cogn. Sci.*, vol. 11, no. 11, pp. 482–488, 2007.

[2] M. E. Shaw, "A Comparison of Individuals and Small Groups in the Rational Solution of Complex Problems," *Am. J. Psychol.*, vol. 44, no. 3, p. 491, 1932.

[3] D. Gigone and R. Hastie, "The Impact of Information on Small Group Choice," vol. 72, no. I, pp. 132–140, 1997.

[4] S. Scherbaum, M. Dshemuchadse, S. Leiberg, and T. Goschke, "Harder than Expected : Increased Conflict in Clearly Disadvantageous Delayed Choices in a Computer Game," *PLoS One*, vol. 8, no. 11, 2013.

[5] I. Lorge and H. Solomon, "Two models of group behavior in the solution of eureka-type problems," *Psychometrika*, vol. 20, no. 2, pp. 139–148, 1955.

[6] R. B. Zajonc, "Social facilitation," *Science (80-.).*, vol. 149, pp. 269–274, 1965.

Do Pitch and Space Share Common Code?

Pulkit Singhal[1], Aditya Agarwala[2], and Priyanka Srivastava[1]

[1]Cognitive Science Lab, International Institute of Information Technology, India
[2]Birla Institute of Technology and Science, Goa
pulkit.singhal@research.iiit.ac.in,f2011433@goa.bits-pilani.ac.in
priyanka.srivastava@iiit.ac.in

1 Introduction

Stimulus-Response Mapping is essential for effective response selection which is important in the course of interaction between perception and action. The perception-action coupling exists in varieties of stimuli, ranging from non-spatial attributes such as color, i.e. Simon effect [1] to spatial correspondence i.e. stimulus and response sharing spatial coding such as pitch and number e.g Spatial-Pitch/Music Association of Response Codes (SPARC/SMARC) and Spatial-Numerical Association of Response Codes (SNARC) respectively [2]. Research shows faster response time when stimulus and response share common coding. Unlike previous studies [2], Beecham et. al [3] has shown reverse and no SPARC effect in addition to SPARC effect, indicating the variability in spatial representation and association with abstract concepts. Based on contradictory findings, we hypothesized that spatial response coding, specifically for SPARC effect, may result from implicit/explicit learning from our day-to-day experience or type of task. To investigate our hypothesis we proposed a series of experiments using pure tones by employing Stimulus-Response Compatibility (SRC) paradigm, similar to Rusconi et al. [2], to measure the implicit and explicit SPARC effect across sighted and visually impaired adults. Unlike Rusconi et al. [2], we aimed to first evaluate the effect of feedback on SPARC effect, which is the focus of the current pilot study. Feedback enables learning and can act as a confound to the implicitness of the phenomenon. Furthermore, considering the individual differences in loudness perception and its interaction with pitch and timbre [4], we examined the difference between equalized and non-equalized loudness on SPARC effect.

2 Methodology

26 musically naive participants (21 male, Mean Age = 25.08) from IIIT Hyderabad volunteered for the experiment. It was five factorial mixed group design: 2 (feedback: with and without feedback, ie. FB) $\times$ 2 (loudness: with and without loudness equalization, i.e. LEQ) as between group and 2 (alignment: vertical and horizontal) $\times$ 2 (arm position: arm and crossarm) $\times$ 2 (congruency: congruent and incongruent) as within group. Participants were randomly assigned to the

LEQ and no-LEQ followed by FB and no-FB conditions. In LEQ, amplitude of the stimuli was subjectively equalized prior to the experiment. In no-LEQ, all the tones shared same amplitude value. Within group conditions were counterbalanced across participants. The task was to compare the pitch of frequency tones E3, F3#, G3#, A3#, D4, E4, F4#, G4# (164.81, 185.00, 207.65, 233.08, 293.66, 329.63, 369.99, and 415.30 Hz) with a fixed reference C4 (261.63 Hz). The participants had to respond whether the target tone was higher or lower in pitch than the reference by pressing the following keys: P/Q for the horizontal alignment; 6/B for the vertical alignment.

3 Results and Discussion

Currently, we are reporting the preliminary data and the results show: 1. trends showing no SPARC effect similar to Beecham et al. [3] indicating heterogeneity in spatial representation and response coding unlike Rusconi et al. [2]. 2. No difference in RT has been observed across feedback and loudness equalization conditions. 3. However, feedback did show a trend of overall better accuracy. However, we made few observations during the course of data evaluation. Rusconi et al. [2] had Cantonese speakers as their participants. Cantonese is a tonal language and research shows that Cantonese speakers have better pitch discrimination abilities [5]. Therefore, it can be argued that the variability in SPARC effect might be due to implicit learning caused by the language or the task performed in everyday experience such as, horizontal SPARC effect as has been observed in musician vs. non-musician.

4 Future Work

Evaluate the individual differences across spatial coding by assessing the cognitive and non-cognitive profile of participants. Explore the role of tonality in language on SPARC effect. Evaluate the type of task by comparing musicians vs. non-musicians and visually impaired vs. sighted population.

References

1. Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing.
2. Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. Cognition, 99(2), 113-129
3. Beecham, R., Reeve, R. A., & Wilson, S. J. (2009). Spatial representations are specific to different domains of knowledge. PLos One, 4(5), e5543.
4. Fletcher, H., & Munson, W. A. (1933). Loudness, Its Definition, Measurement and Calculation*. Bell System Technical Journal, 12(4), 377-430.
5. Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. PLoS One, 8(4), e60676.

Recent response conflict modulates early distractor processing

Michael Sprengel

Freie Universität Berlin
Habelschwerdter Allee 45, 14195 Berlin, Germany

Markus Hofmann

Freie Universität Berlin
Habelschwerdter Allee 45, 14195 Berlin, Germany

Mike Wendt

Helmut Schmidt University / University of the Federal Armed Forces Hamburg
Holstenhofweg 85, 22043 Hamburg, Germany

Aquiles Luna-Rodriguez

Helmut Schmidt University / University of the Federal Armed Forces Hamburg
Holstenhofweg 85, 22043 Hamburg, Germany

Sascha Tamm

Freie Universität Berlin
Habelschwerdter Allee 45, 14195 Berlin, Germany

Thomas Jacobsen

Helmut Schmidt University / University of the Federal Armed Forces Hamburg
Holstenhofweg 85, 22043 Hamburg, Germany

Arthur M. Jacobs

Freie Universität Berlin, D. I. N. E.
Habelschwerdter Allee 45, 14195 Berlin, Germany

michael.sprengel@fu-berlin.de

Abstract. Flooded with a myriad of stimuli impinging on our retinae at almost every waking moment, we must have formidable sluices to tame the incoming tide – or drown in hopeless confusion. As theorized about the location and (cognitive) control of these sluice gates, recent ERP-findings indicate an early perceptual one, whose permeability depends on the *average* utility of information passing through. The remarkable ability of humans to adapt and even switch task sets seemingly instantaneously suggests that, beyond the reported *block-wise* effect, trial-by-trial fluctuations of selectivity should be observable. To investigate this possibility, we modified the authors' temporal flanker paradigm to include 50% of both congruent and incongruent trials in each block. As expected, the simultaneous ERP-recordings featured a higher posterior visual N1 in response to flankers following a congruent trial. Taken to reflect more intensive processing, this explains the distractors' stronger influence on behavior in this condition, as evidenced by a higher interference effect than after an incongruent trial. This typical sequential interference modulation was mirrored by a target-related fronto-central N2 believed to originate from medial prefrontal Cortex (mPFC). Among models of mPFC-activity, the Conflict Monitoring Theory (CMT) provides a comprehensive interpretation of the reported patterns. It even explains the additionally observed modulation of the flanker-N1 by the penultimate trial's congruency – as a further manifestation of the same response conflict-induced perceptual selectivity adjustments that the direct sequential N1-modulation represents. In stark contrast, the predicted response-outcome (PRO) model fails to fit any of these apparent attentional modulations of early distractor processing.

Keywords: Response conflict, selective attention, visual N1

Emotional Effects on Time Estimates during Intervals up to 5s

Anna Katharina Trapp[1], and Manfred Thüring[1]

[1]Technische Universität Berlin, Chair of Cognitive Psychology and Cognitive Ergonomics, Berlin, Germany
{anna.k.trapp,manfred.thuering}@tu-berlin.de

Abstract. Many studies have shown an effect of arousal and valence on time estimates for short duration (<2 s), while cognitive models of prospective timing only predict an effect of arousal. For durations up to 6 s the evidence is not as clear. Therefore, we aimed to replicate the effect of emotions for short durations and to obtain deeper understanding of emotional effects during longer durations. In a study we induced arousal (via volume of sound) and valence (negative vs. neutral) by aural stimuli during durations between 1.1 and 5.0 s. We measured over- and underestimation in a temporal reproduction task and pupil dilation as an indicator for arousal. Higher volume led to weaker underestimations at all durations, while negative sounds were more underestimated than neutral sounds. The results replicate the prolonging effect of emotions on time estimates up to 5s indicating an effect of arousal and valence.

Keywords: prospective time perception, emotions, clock speed hypothesis, auditory mode

1 Introduction

A prolonging effect of emotions on duration estimates for attended intervals shorter than 2 s has been frequently shown with different timing tasks and different induction methods [1]. This effect is commonly attributed to the clock speed hypothesis which ascribes the effect to an increased arousal level during emotional episodes. However, some studies report not only an effect of induced arousal level but also an effect of the valence of stimuli [2,3] for durations under 2 s. When looking at intervals of exactly 2 s, findings are not as clear. Noulhiane et al. [3] focused on the effect of high arousing versus low arousing sounds with either negative or positive valence on a reproduction and a verbal timing task. They report longer estimates for low arousing sounds compared to high arousing sounds at durations of 2 s for both positive and negative stimuli. These findings are contrary to the clock speed hypothesis. While Noulhiane et al. [3] used aural stimuli, Angrilli et al. [3] conducted a similar study with visual stimuli. At 2 s durations, they replicated the arousal effect predicted by the clock speed hypothesis. As an explanation for the reversed effect as reported by Noulhiane et al. [3], the authors argue that it could be a modality specific effect. Looking at intervals lasting longer than 2 s, evidence suggests that the effect of emotions disappears [2,3,4]. Even though the clock speed hypothesis is not limited to a specific duration, it seems like its predicted arousal effect is short-lived. Based on the discussed literature the following issue is addressed in our study: How do arousal and valence affect time estimates of durations ranging from under 2 s to 5 s in the auditory mode? To address this issue, arousal and valence of aural stimuli were systematically varied. By concentrating on this range of durations we focus on both durations for which the prolonging effect seems reliable respectively unclear.

2 Method

An experiment with N=20 participants (11 male, M_{age}=25.1 years, SD_{age}=3.3) was conducted with a 4x3x2 within subject design consisting of the factors duration (1.1 s, 2.4 s, 3.8 s, 5.0 s), volume (M_{low}=41 dBA, M_{medium}=61 dBA, M_{loud}=70 dBA) and valence of stimuli (negative, neutral). The factor volume was chosen as an arousal manipulation because continuous noise as in sounds has been argued to increase arousal level [5]. The three volume levels were applied to 12 sounds taken from the IADS database [6]. They were chosen based on arousal and valence level [6]. Respectively 6 of them had either neutral or negative valence, $t(19)=-10.8$, $p<.001$. However, the two valence categories also differed with respect to arousal, $t(19)=4.7$, $p<.001$. Reproduced time estimates were the main dependent variable. Pupil dilation served as manipulation check because it is an indicator for changes in arousal level. Participants were instructed not to count. During the test phase 180 trials were presented (4x3x2 conditions x6 sounds + 36 distractor trials). Each trial consisted of a 2 s delay to measure the baseline of pupil diameter, the presentation of the stimulus, a reproduction task and the question whether the participant had counted. Pupil diameter was measured during the whole trial. For data analysis, all distractor trials and “counted trials" were excluded. The pupil dilation for each trial was calculated by subtracting the averag pupil diameter of the left eye during baseline from the averaged pupil diameter during stimulus presentation. Regarding time estimates, perceived time ratios (PTR) were calculated by dividing the reproduced by the presentation duration. PTRs allow to compare estimates concerning strength of under- and overestimations independently from different intervals. Effect size was computed by generalized eta-square.

3 Results

In almost all conditions, mean PTRs were smaller than 1 thus indicating underestimations. An ANOVA on PTRs showed main effects for duration, $F(1.16, 21.99)=54.2$, $p<.001$, $ges=.286$, volume, $F(2,38)=20.8$, $p<.001$, $ges=.025$, and valence, $F(1,19)=15.3$, $p<.001$, $ges=.008$. Sounds showed an approximately linear decrease of PTRs from durations of 1.1 s to 5.0 s. PTRs increased with rising volume. Negative sounds showed higher PTRs than neutral ones. The ANOVA also showed an interaction between duration and volume, $F(2.64, 50.21)=4.9$, $p<.01$, $ges=.012$. Closer inspection revealed that the effect of volume on PTRs was strongest for 1.1 s durations. The effect was smaller but still present at 3.8 and 5.0 s. Regarding pupil dilation, an ANOVA showed a main effect of duration, $F(1.55, 29.44)=6.7$, $p<.001$, $ges=.025$ and a main effect of volume, $F(2, 38)=22.3$, $p<.001$, $ges=.052$. Pupil diameter increased more when hearing sounds for a longer durations compared to a shorter duration and it increased more when the volume was higher. The ANOVA also revealed a marginally significant interaction of duration and volume, $F(6, 144)=2.2$, $p=.052$, $ges=.008$. The increase of pupil dilation over the four durations was marginally stronger for high volume compared to low volume.

4 Discussion

This study investigated how arousal and valence induced by aural stimuli affected duration estimates between 1.1 and 5.0 s. Measured pupil dilation indicated that arousal induction via volume of sound worked well. As predicted by the clock speed hypothesis, increased arousal level led to longer time estimates for all durations. Compared to other findings, the presented data does not replicate the disappearance of an arousal effect for durations longer than 2 s. A reason for this result may be the more intense and less avoidable induction method via volume. However, the arousal effect was weaker for longer intervals. Additionally, the results did not implicate a turning of the arousal effect at 2.4 s for aural stimuli. Regarding valence, negative and neutral sounds also differed with respect to arousal level. Therefore, the finding that higher estimates were given for negative compared to neutral stimuli can not be clearly ascribed to a change in valence. However, pupil dilation did not differ between valence groups indicating that the difference on arousal level might be rather small. Thus, we interpret the data as weak evidence for a valence effect but more strongly as a prolonging effect of emotions on time estimates. Concerning duration, all four durations were underestimated which was stronger for longer sounds indicating either a decrease of arousal level or additional attentional processes during longer intervals. The effect of duration on pupil dilation can be explained by the data analysis strategy. Because pupil diameter following the sound onset showed a continuous increase after a short delay, during short durations the maximum plateau was often not reached within the presentation duration lowering the averaged value. Hence, the duration effect on pupil dilation is probably not related to arousal differences between different durations. To summarize, the results replicate the prolonging effect of emotions on prospective time estimates emphasizing the importance of arousal but giving weak evidence for a valence effect as well. In line with the clock speed hypothesis but contrary to other literature, the arousal effect did not vanish after 2 s. However, it was weaker for longer durations. We ascribe this finding to a more intense arousal induction which should be further investigated.

5 References

1. Droit-Volet, S., Gil, S.: The time-emotion paradox. Philosophical Transactions of the Royal Society B 364, 1943–1953 (2009) http://doi.org/10.1098/rstb.2009.0013
2. Angrilli, A., Cherubini, P., Pavese, A., Manfredini, S.: The influence of affective factors on time perception. Perception & Psychophysics 59(6), 972–982 (1997) http://doi.org/10.3758/BF03205512
3. Noulhiane, M., Mella, N., Samson, S., Ragot, R., Pouthas, V.: How emotional auditory stimuli modulate time perception. Emotion 7(4), 697–704 (2007) http://doi.org/10.1037/1528-3542.7.4.697
4. Nather, F.C., Bueno, J.L.O., Bigand, E., Droit-Volet, S.: Time changes with the embodiment of another's body posture. PLoS ONE 6(5), e19818 (2011) http://doi.org/10.1371/journal.pone.0019818
5. Hockey, G.R.J.: Effects of noise on human efficiency and some individual differences. Journal of Sound and Vibration 20(3), 299–304 (1972) http://doi.org/10.1016/0022-460X(72)90610-4
6. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL (2008)

Predicting patterns in navigator-driven placement of landmarks for future wayfinding with Space Syntax

Rul von Stülpnagel[1] and Christoph Hölscher[2]

[1] Freiburg University, Germany
rul.von.stuelpnagel@cognition.uni-freiburg.de;
[2] ETH Zürich, Switzerland
christoph.hoelscher@gess.ethz.ch

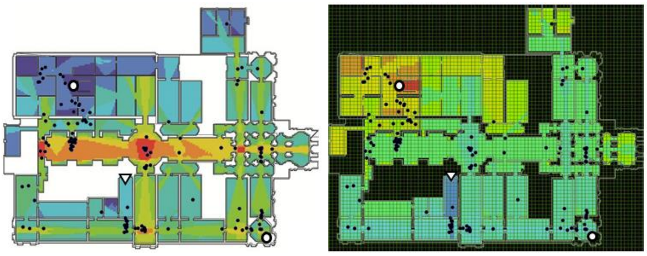
Abstract. We investigated how navigator-driven landmark placement for the facilitation of future wayfinding tasks is affected by the spatial configuration of the environment. A space syntax analysis suggests that participants placed landmarks at more integrated locations of a virtual environment in general. More interestingly, participants' landmark placement apparently represents a trade-off between a desired proximity to the designated goal location and maximized visibility of the landmark. This trade-off can be described as a function of the visibility steps' relative weight in regard to the goal location and global integration.

Keywords: Space syntax, landmarks, wayfinding, structural salience

1 Description

Research on the role of landmarks for human navigation has been mostly concerned with landmarks available in the surrounding environment. A different approach consists of analyzing how and where landmarks are set up by individuals to facilitate current and future orientation [1]. In a recent study, we found strong consistencies in landmark placement across participants. The paper at hand extends the analyses of [2], Study 2, with the aim to link the observed patterns in landmark placement to the spatial properties of the environment by applying the space syntax methodology [3]. Space syntax properties are predictors of both wayfinding behavior [3] and landmark selection [4]. We expected landmark placement to be biased towards more integrated spaces in general. We also hypothesized that landmark placement thought to signpost a goal location is a trade-off between the maximization of a landmark's integration within the building's layout and a sufficient proximity to a goal location.

All details concerning study design and procedure are described in [2], Study 2. We performed a visibility graph analysis of the building's layout based on a 1×1m grid to capture its spatial properties. We counted the number of landmarks placed in any given cell by participants in the two experimental conditions featuring navigator-driven landmark placement (N = 33), with a total number of landmarks of N = 157. A Spearman correlation between a cell's visual integration and number of placed landmarks indicated that more integrated cells were indeed preferred, r = .11, p > .000. In particular, the core of the building attracted a large proportion of landmarks. Next, we quantified the proximity to a goal location by running a visibility step analyses, separately for all three goal locations. Visibility steps (VE) closer to goal locations did not receive more landmarks per se. We weighted a visibility step's relevance by computing the ratio of its visually most integrated cell (VMIC) and its proportion of the whole environment (PE). VMIC/PE proved to be a strong predictor of the density of landmarks placed within a visibility step for all three goal locations (LM density, see Fig.1).



VS	PE	VMIC	VMIC/PE	LM Density
#1	.02	3.99	200	.30
#2	.09	5.11	57	.09
#3	.26	5.68	22	.06
#4	.32	5.55	17	.03
#5	.11	3.99	36	.03
#6	.09	2.99	33	.03
#7	.07	2.35	34	.06
#8	.04	1.92	48	.04

Fig. 1. Left: Visual integration of the test environment, with blue to red indicating increasing integration values. Center: Visibility steps starting from a goal location indicated by the white triangle is presented, with blue to red indicating the increasing visibility step number. (Two other goal locations used in the study are marked with white circles). Black dots indicate individually placed landmarks. Right: Table illustrating the relation of a visibility step's relative weight and landmark placement density for the goal location indicated by the white triangle.

We conclude that navigator-driven landmark placement used to sign-post a goal location can be predicted by Space Syntax. Participants balanced the need to establish a connection of goal location and landmark on the one hand, but to increase the landmark's usefulness by maximizing its visibility and integration on the other hand. In other words, humans are highly sensitive to the spatial properties of their environment. Our findings may help to develop signage solutions for critical locations optimized for human wayfinding needs based on the analysis of a building's spatial properties.

2 References

[1] D. Cliburn, T. Winlock, S. Rilea, and M. Van Donsel, "Dynamic landmark placement as a navigation aid in virtual worlds," in *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, 2007, pp. 211–214.

[2] R. Von Stülpnagel, S. Kuliga, S. J. Büchner, and C. Hölscher, "Supra-individual consistencies in navigator-driven landmark placement for spatial learning," in *Proceedings of the 36th annual conference of the cognitive science society. Cognitive Science Society, Austin, TX*, 2014, pp. 1706–1711.

[3] R. Conroy-Dalton, "Spatial Navigation in Immersive Virtual Environments," University of London, 2001.

[4] R. von Stülpnagel and J. Frankenstein, "Configurational salience of landmarks: an analysis of sketch maps using Space Syntax," *Cogn. Process.*, vol. 16, no. S1, pp. 437–441, Sep. 2015.

3D-Shape-Perception studied exemplarily with Tetrahedron and Icosahedron as prototypes of the polarities Sharp versus Round

Iris Sauerbrei[1], Jörg Troja[2], Erich Lehner[3]

[1]Techincal University Vienna, Hohenfriedbergstraße 20, 10829 Berlin, Germany
[2]University Koblenz-Landau, Fortstraße 7, 76829 Landau, Germany
[3]Technical University TU Wien, Austria, Karlsplatz 13/251, 1040 Wien, Austria
irisesthersauerbrei@hotmail.com, trojan @uni-landau.de, erich.lehner@tuwien.ac.at

Abstract. Platonic solids as walk-in-models explore perception of three-dimensional shapes in dependence on the approach of round versus sharp, which was examined by research studies in psychology of perception and neuroscience in surveys by e.g. Richter and Hentsch, Bar and Neta, Nanda and Pati. Among the five Platonic solids, tetrahedron and icosahedron provide the clearest representation of the qualities sharp and round, respectively. By a questionnaire preferences between the two walk-in-models (height 2.2m) were tested. Preliminary analyses of the collected data provide first evidence that the shapes of basic three-dimensional geometrical objects are associated with specific patterns of attributes and association.

Keywords: 3D Shapes, Geometrical features, Space Perception, Round versus Sharp, Walk-in models, Architecture, Spatial cognition.

1 Introduction

Geometrical features characterize three-dimensional formations and affect the perception of them. Therefore, particular geometric features should correspond to specific, identifiable perceptions.

2 Motivation / Research Target

To enable conscious choices for spatial arrangements in design processes for the built environment, it is necessary to identify characteristic attributes of three-dimensional shapes. An architect has access to criteria of functionality and structural analysis for assessing a proposed design, but those to validate the quality of shapes are absent. Main research target of this study is to initiate a process to close this gap.

3 Experimental Component

3.1 Geometry

The five platonic solids [1] are the geometrical base for all other three-dimensional shapes [2], among which tetrahedron and icosahedron provide the clearest representation of the qualities sharp and round.

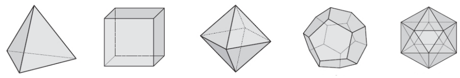


Fig. 1. The five Platonic Solids

3.2 Precedent Surveys

This pair of attributes has already been examined in various surveys in psychology of perception and in neuroscience. In the 1930s, psychologist Wolfgang Köhler defined a pair of two-dimensional shape archetypes - "Takete" and "Maluma"- embodying the qualities round versus sharp [3]. A study of Peter G. Richter and Norbert Hentsch related them to the designing process for objects of daily use and identified several associated attributes and sensations [4].

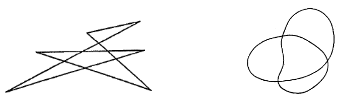


Fig. 2. Takete and Maluma according to W. Köhler

Neuroscientists Bar and Neta showed that patterns and objects of daily use with soft, round shapes are preferred in comparison to those with sharp and pointed profiles.[5] Further research demonstrated the valid transfer to large-scale surroundings like architecture and landscape by walk-through videos and photographs.[6] [7].

3.3 Experiment

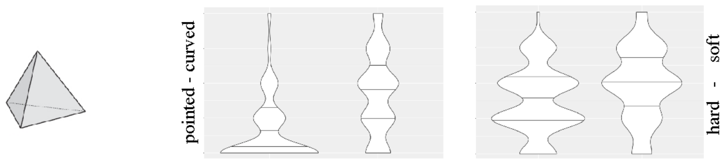
To verify a transfer of outcomes of the existing surveys to the perception of three-dimensional shapes, walk-in models (total height 2.2m) of tetrahedron and icosahedron were set up in a public park in Frankfurt am Main, Germany, to test preferences and to collect associations of park visitors by questionnaire.
In summer 2015, we assembled the tetrahedron eight times, the icosahedron seven times. In total 288 participants took part in the study; 116 rated the tetrahedron, 172 rated the icosahedron.

Fig. 3. Walk-in models of tetrahedron and of icosahedron
with participants in Niddapark, Frankfurt/ Main 2015

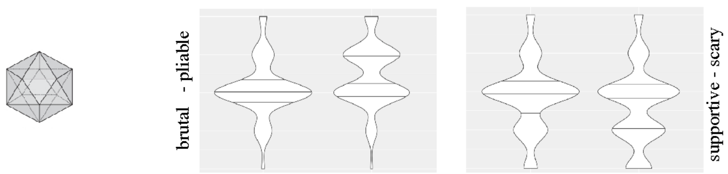
4 Findings

The findings provide first evidence that the shapes of basic three-dimensional geometrical objects are associated with specific patterns of attributes and associations.

Using Wilcoxon Rank-Sum tests, preliminary analyses of the collected data to the preference ratings show that participants rated the tetrahedron as significantly narrower, more pointed, sharper, more familiar and harder than the icosahedron; the icosahedron was perceived as significantly more pliable, more endearing, more supportive, and more inspiring than the tetrahedron.



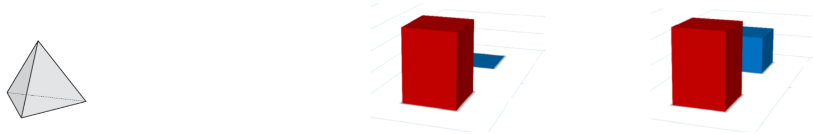
The Tetrahedron is perceived as more pointed and harder.



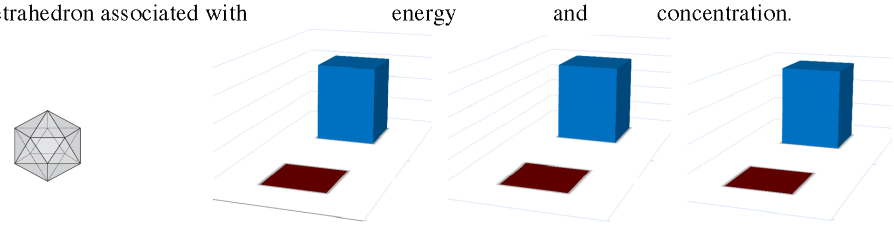
The Icosahedron is perceived as more pliable and supportive.

Fig. 3. Diagrams of Wilcoxon Rank-Sum test with continuity correction

By sorting and counting the associations given by test persons it becomes apparent, that the tetrahedron is perceived as being simple and vertical in direction in regards to its geometry, and generally stands for energy and concentration. Therefore it represents qualities like focusing and realization. Different from this, the icosahedron is perceived as complex and spreading in circular direction in regards to its geometry, and generally stands for integration, community, future and emotions of happiness. Therefore it represents qualities like communication and atmosphere.



Tetrahedron associated with energy and concentration.



Icosahedron associated with integration / community, future and emotions like happiness.

Fig. 4. Excel- Diagrams showing quantities of associations

4 Conclusions

The fact that specific shapes may elicit distinct patterns of perceptions is relevant for all fields of design.

Future studies should examine similarities and differences between the perceptions of two- and three-dimensional shapes and study attributes of shapes that are more complex. Additionally, attributes for the three other basic bodies - Octahedron, Cube, and Dodecahedron – need to be searched for.

References

1. Eukleides of Alexandreia (Euklid): 'Elements' in Books X and XIII (Mathematical evidence of Theaitetos)
2. Ziegler R.: `Platonische Körper- Verwandtschaften, Metamorphosen, Umstülpungen', Verlag am Goetheanum, Dornach, 2008, pp. 45 + 48
3. Köhler W.: `Psychologische Probleme', Verlag Julius Springer, Berlin, 1933, p. 153
4. Richter P.G., Hentsch N.: `Takete und Maluma, Eine Untersuchung zur Herkunft von (ikonischen) Vorstellungen in frühen Phasen des Produktentwurfes', TU Dresden, ca. 2005
5. Bar M., Neta M.: Visual elements of subjected preference modulate amygdala activation', 2007, http://www.sciencedirect.com/science/article/pii/S0028393207001042
6. Hobbs H.: Description of Diploma Thesis (Interior space simulations by Walk-Through Videos) in Conference Catalogue ANFA 2014, p. 52
   Pati Dr. D.: Description of Survey (MRT scans providing photographs of landscapes, buildings, interiors, objects within the machine) in Conference Catalogue ANFA 2014, p. 84
7. Vartanian et. al.: `Impact of contour on aesthetic judgments and approach-avoidance decisions in architecture`, 2013, PNAS Early Edition, www.pnas.org/cgi/doi/10.1073/pnas.1301227110, p.1

Author BiOS

Dipl. Ing. Iris Sauerbrei is a practicing architect who participated at designs and realizations e.g. of public buildings such as Jewish Museum Berlin 1997, Cultural Centre Abu Dhabi 2008, Station Buildings Doha Qatar 2011. For her advanced staged PHD she chose the subject space perception and attended the conferences ANFA 2014 and ICSC ROME (6th International Conference on Spatial Cognition).

Jun.-Prof. Dr. Jörg Trojan is a university professor at the University Koblenz-Landau, Campus Landau, at the Department of Psychology, where he teaches classes on biological psychology and human-environment interaction. His research focuses on how we perceive our own bodies and the surrounding space.

Prof. Dr. Ing. Erich Lehner is a university professor at the technical university Vienna, where he teaches since 1982. Until today, he initiated and carried out numerous research projects, lectures and publications at home and abroad. He is the PHD supervisor for the presented study.

Representation of Wayfinding and Perception Abilities in Agent-Based Models

Erik Andresen

Chair of Fire Safety and Pedestrian Traffic,
School of Architecture and Civil Engineering, University of Wuppertal
{andresen}@uni-wuppertal.de

Abstract. The design, management and control of pedestrian traffic systems are essential tasks to retain mobility facing increasing traffic volume world wide. To be able to forecast traffic distributions it is important to understand the traffic participants' wayfinding behaviour as well as the physical movement patterns. My work as a PhD student comprises the creation of a symbolic agent-based computer model representing selected procedures of the human wayfinding process. Particularly, I focus on the cognitive map and its mode of action. The model is connected to a simulation framework modeling the locomotive behaviour of pedestrians.

Keywords: Wayfinding, Cognitive Map, Model, Pedestrians

1 Introduction & Motivation

The control or rather the management of traffic is obviously an important task, particularly in regions where the amount of traffic participants is still increasing. Pedestrian traffic flows on street networks, in public buildings, train stations or shopping malls, etc. need to be optimized and congestion to be avoided regarding situations of daily traffic and emergency cases. To predict the progress of traffic systems, i.a. computer simulations based on microscopic agent-based models are utilized.

For realistic forecasts of the distribution of traffic participants or rather flows and congestion, the participants' wayfinding behaviour needs to be regarded and examined as well as the locomotion. However, in many simulation frameworks exit choice decisions and wayfinding are only roughly represented by shortest path algorithms or similar minimum effort calculations. Within the context of the PhD thesis I aim to close this gap by introducing a computer model considering both the locomotive actions and the solving of wayfinding tasks. Therefor I would like to figure out the strategies that are used to solve a wayfinding task in a specific environment or rather in specific circumstances. Furthermore, I aim to investigate the underlying processes and tools and how they can be modeled adequately so that the approaches can be combined with locomotion models.

Already done work comprises symbolic modeling approaches representing selected human wayfinding abilities. They are embedded into the pedestrian simulation framework *JuPedSim* [1]. The basic ideas of these modeling approaches are discussed in the following sections.

Beside the approaches which represent wayfinding abilities, the presented model provides agents with perceptional abilities and enables them to make reasonable route choice decisions based on both the current visible environment and their memorized spatial information.

2 Modeling incomplete, inaccurate knowledge (the cognitive map)

The wayfinding procedure is a non-trivial process comprising various tools and strategies. The cognitive map plays one of, if not even, the most important role in the solving of a wayfinding task. It describes the mental representation of a wayfinder's large-scale spatial environment including landmarks (salient remembered objects) and their spatial relationships to each other [2, 4]. In principal, the wayfinding model presented in this paper is a representation of the cognitive map and its characteristics.

Creating a model representing the setup and functionality of the cognitive map I particularly take the following map's properties into account. Most importantly, in the majority of cases the information in the map is inaccurate, $fuzzy$, distorted, partially or may be even completely wrong [4]. Thus, I propose to describe the cognitive map to be a construction of ellipse items (see Fig. 1) [3].

The ellipses represent the inaccurate estimation of a landmark's location. Depending on the knowledge degree about a landmark's location size, shape and position of the ellipse is set. The more inexact the memory about the location the greater the ellipse and its shifting from the real position [3]. Still vague information about the spatial relationships between landmarks (for example: "Landmark A lies beyond Landmark B") can be extracted from the ellipses' locations. However, exact metrical relations will not be evaluated since the exact positions of the landmarks within the ellipses are not clear.

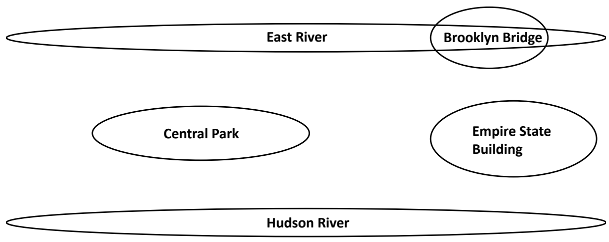
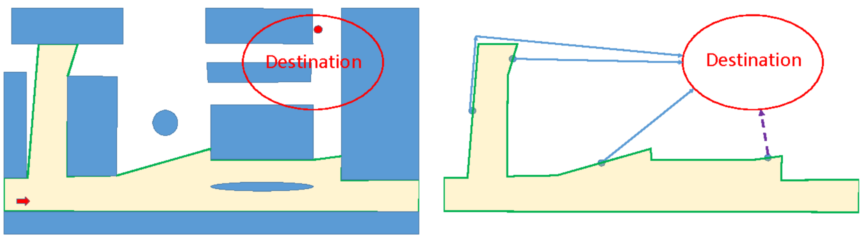


Fig. 1. A conceivable cognitive map of New York City (Manhattan). The figure depicts how the map's landmarks are described in the model (by ellipses).

To demonstrate how the agent uses the information in his cognitive map I make use of the following example. I consider a simulated pedestrian to be located in the fictional environment shown in Fig. 2(a). His position is marked by the red arrow in the left lower corner. He tries to reach the destination in the upper right corner which could be his main target or a landmark serving as a subgoal. The ellipse represents the destination's position in the agent's cognitive map. The red dot marks the real destination's position. The color coded area depicts the area which the pedestrian has visual access to. It has been determined by using a visibility graph. The edges of the visibility polygon can be split up into either edges that border on objects and crossings. In the example scenario the agent has to choose between various possibilities (crossings) to get closer to his destination. As a representation of the decision making to select a proper proceeding way the following steps will be executed. First the visibility polygon and the target ellipse are regarded without exception (see Fig. 2(b)). The next step comprises the determination of the shortest paths between every crossing (in Fig. 2(b) represented by a blue dot at the center of the crossing) and the boundaries of the target ellipse [3]. The created paths must not intersect with the visibility area or rather with visible walls or other crossings. Furthermore the length of the paths are calculated and compared with each other [3]. The agent assumes that the crossing related to the path possessing the shortest path length (the shortest one of all shortest paths) is the best choice to come closer to his target [3].

While heading towards the crossing which is presumed to be the most appropriate one the visibility area certainly changes or rather grants visual access to another part of the geometry. With the change of the visibility area also the crossings' positions will shift and new crossings may emerge. In consequence, I enable the agent to react on the changes and grant him the capability to perceive and evaluate his visible surroundings in every desired time step of the simulation.



(a) (b)

Fig. 2. a) Overview of a fictional spatial environment. Visible area of the agent is color coded. The (only) memory in the agent's cognitive map is depicted by the ellipse. b) Visible area and target ellipse separated from the geometry. The shortest paths from every crossing of the visible area to the ellipses are marked. The shortest one of all paths is shown dashed.

3 Summary & Conclusion

The presented model provides a simulated agent with a cognitive map incorporating approximate knowledge of landmarks' locations. The map can be modified by changing size, position and shape of the ellipses which represent the map's landmarks [3]. Thus, in a scenario including multiple agents, it is possible to provide every single agent with an individual distinguishing knowledge degree (represented by each cognitive map). In addition, the agents have the abilities to perceive and evaluate their visible surroundings and select a proper crossing that brings them (at least) closer to their (sub-)targets [3]. The model can be used to simulate the wayfinding behaviour in indoor facilities and outdoor areas.

As the presented model is connected to pedestrian locomotion models the consequences of specific set-ups of landmarks in the cognitive maps, for example the walked path (trajectories) of the agents, can be regarded and investigated. Based on the agents' trajectories the pedestrians' distributions within an area or building can be scrutinized. In addition, with the help of the model the legibility of a facility depending on various knowledge degrees can be tested. Particularly, designers or controllers of facilities or events can use the combined model to evaluate the influence of changing the position or amount of very salient landmarks to the distribution of the agents.

As has been shown by [5] the wayfinding process comprises additional tools and strategies beyond the here mentioned ones. Thus, I'm interested in the investigation and modeling of further cognitive processes that are involved in the wayfinding procedure. My long-term goal comprises the creation of a framework combining the modeling of all tools and strategies playing a major role in wayfinding tasks.

Acknowledgments. This research is founded by the Deutsche Forschungsgemeinschaft (DFG) contract No. GZ: SE 17894-1.

References

1. Chraibi, M., Kemloh Wagoum, A.U.; Zhang, J.: JuPedSim: An open framework for simulating and analyzing the dynamics of pedestrians. In 3rd Conference of Transportation Research Group of India (in press).
2. Tolman, Edward C.: Cognitive maps in rats and men. In: Psychological Review 55 (4), pp. 189–208. DOI: 10.1037/h0061626 (1948).
3. Andresen, Erik; Haensel, David; Chraibi, Mohcine; Seyfried, Armin: Wayfinding And Cognitive Maps For Pedestrian Models. In: Proceedings of 11th International Conference on Traffic and Granular Flow 2015 (in press)
4. Golledge, Reginald G.: Human Wayfinding and Cognitive Maps. In: Reginald G. Golledge (Hg.): Wayfinding behavior. Cognitive mapping and other spatial processes. Baltimore: Johns Hopkins University Press, pp. 6–45 (1999)
5. Wiener, Jan M.; Büchner, Simon J.; Hölscher, Christoph: Taxonomy of Human Wayfinding Tasks: A Knowledge-Based Approach. In: Spatial Cognition & Computation 9 (2), pp. 152–165. DOI: 10.1080/13875860902906496 (2009)

Inductive Learning of Categories
Between Cognitive Modeling and Machine Learning

Christina Zeller

Cognitive Systems Group, University of Bamberg
An der Weberei 5, 96045 Bamberg, Germany
{Christina.Zeller}@uni-bamberg.de

Abstract. Categorization is an essential cognitive skill. While the construction of categories from examples is the focus of machine learning, most theories in cognitive science address the representation of categories and categorization. The main goal of my thesis is to renew the connection between cognitive modeling and machine learning. Based on recent empirical work, I am interested in rule-based learning. I plan to conduct psychological experiments to uncover algorithmic principles of category learning and I plan to realize a cognitive model in Prolog which shall cover several aspects of category learning.

Keywords: category learning, cognitive modeling, machine learning

Categorization is an essential cognitive skill, which provides us, for example, to make prospects about objects of a category and to communicate using a specific word for a group of objects. Different theories try to describe how humans categorize, mainly the rule-based, the prototype, the exemplar and the decision-boundary theory (Kruschke, 2008). In addition, hybrid theories combine these approaches and augment them with ideas declaring their interaction. Hybrid theories often include the rule-based approach to explain categorization using logical rules. Rule-based approaches were researched a lot in the middle and the end of the 20th century (cf. Bruner, Goodnow, & Austin, 1956; Hunt, Marin, & Stone, 1966; Unger & Wysotzki, 1981). However, for a long time they were criticized for having shortcomings explaining typicality effects, which means that typical members of a category are treated different than non-typical members. Nevertheless, it was shown recently that typicality effects can be explained by a rule-based theory (Lafond, Lacouture, & Cohen, 2009).

Most cognitive approaches of categorization focus on representation of categories and categorization. That is, they do not describe the underlying learning algorithms. In machine learning research these underlying learning algorithms are essential (Mitchell, 1997). Early machine learning algorithms were often inspired by psychological findings of categorization, but over the years the connection got weaker (Langley, 2016). **With my research I want to renew the connection between cognitive modeling and machine learning by modeling how humans learn to categorize and whether the thereby modeled humans' strategies can be applied as resource for developing machine learning algorithms.**

As a first step, we started to analyze data from a categorization experiment realized by Lafond et al. (2009) where we focused on learning categories. In this experiment, five participants learned to categorize images of 3D rendered lamps with a trial-by-trial feedback. The resulting categorization process was modeled with individual decision-trees. It was possible to predict the structure of these decision-trees with a measure we call *igain* (Zeller & Schmid, accepted). Furthermore, we tried to model the answering behavior of the five participants during the learning phase with the incremental decision-tree algorithm CAL 2 (Unger & Wysotzki, 1981). The matching of answers during the learning phase from the five participants and CAL 2 was between 61 % and 88 %. Therefore, we concluded that the five participants did not use a purely incremental learning algorithm as implemented in CAL 2.

Additionally, I analyzed the time behavior during the learning. In the experiment participants saw a lamp, categorized it, got feedback on their categorization and then could ask for the next example on their own pace. The time for asking for the next lamp was higher after a negative feedback. That could indicate that participants changed their categorization hypotheses after an incorrect categorization, which is in line with earlier theories (cf. Bruner et al., 1956).

To come up with a model capturing empirically observable strategies of human category learning, several aspects have to be investigated. Mainly, open questions concern (a) the representation language for learned hypotheses (i.e. categories), (b) feature selection strategies, and (c) the learning strategy. Rule-based learning implies that the *representation* can have the format of sets of rules or a compact representation as a decision-tree. Rules might be restricted to conjunctions of a fixed low number of features. In its most extreme, humans try to represent a category by a single attribute (Goede & Klix, 1972). Alternatively, rules might be composed of conjunctions or disjunctions of attributes. Furthermore, negation of feature attributes should be considered. *Feature selection* is eventually influenced by observed co-occurrences between feature attributes and categories, but might be also influenced by perceptive salience (cf. comment in Zeller and Schmid regarding utility values in Lamberts, 2000). The *learning strategy* can be assumed to be incremental rather than batch. Open questions concern whether one representation is systematically constructed over the whole learning process or whether meta-strategies are involved. Such meta-strategies could influence the deletion of hypotheses, or switching between different representation formats (Goede & Klix, 1972). For example, humans might first try to represent a category by a single feature and only switch to more complex representations after failure. Furthermore, in a categorization task including two categories (e.g. A vs. ¬A) context might influence whether a representation for Category A or the negation (¬A) is constructed. For example, if many negative examples are presented at the beginning of the learning phase, the learner might be biased towards ¬A.

These questions shall be answered with psychological experiments to uncover algorithmic principles of category learning. The resulting algorithmic principles are planned to be realized as a cognitive model in Prolog which can incrementally learn propositional rules represented as feature vectors. Additionally, humans are able to understand and learn relational and recursive rules (cf. Besold, Muggleton, Schmid, Tamaddoni-Nezhad, & Zeller, accepted), for example, the category of a father relation and the category of an ancestor relation. These categories, depending on relational and recursive rules, are the focus of inductive logic programming which is a part of machine learning research (cf. Mitchell, 1997). However, inductive logic programming algorithms usually learn from a complete example set instead learning incrementally. Therefore, I plan to investigate human, incremental, relational and recursive rule learning of categories and expand my cognitive model according to my findings. This shall lead to new insights in the category learning research as well as inspire new machine learning algorithms and therefore renewing the connection between cognitive modeling and machine learning.

References

Besold, T., Muggleton, S., Schmid, U., Tamaddoni-Nezhad, A., & Zeller, C. (accepted). How does predicate invention affect human comprehensibility? In *26th International Conference on Inductive Logic Programming (London, England, Sept. 2016).*

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking: With an appendix on language by Roger W. Brown.* New York, NY: Wiley.

Goede, K., & Klix, F. (1972). Lernabhängige Strategien der Merkmalsgewinnung und der Klassenbildung beim Menschen (Human learning-dependend strategies of feature selection and categorization). In F. Klix, W. Krause, & H. Sydow (Eds.), *Kybernetik-Forschung (Cybernetics Research)* (Vol. 1, pp. 151–212). Berlin, Germany: Deutscher Verlag der Wissenschaften.

Hunt, E. B., Marin, J., & Stone, P. J. (1966). *Experiments in induction.* New York, NY: Academic Press.

Kruschke, J. K. (2008). Models of categorization. In R. Sun (Ed.), *Computational psychology* (pp. 267–301). Cambridge University Press.

Lafond, D., Lacouture, Y., & Cohen, A. L. (2009). Decision-tree models of categorization response times, choice proportions, and typicality judgments. *Psychological Review*, *116*, 833–855.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.

Langley, P. W. (2016). The central role of cognition in learning. *Advances in Cognitive Systems*, *4*, 3–12.

Mitchell, T. M. (1997). *Machine learning.* McGraw-Hill.

Unger, S., & Wysotzki, F. (1981). *Lernfähige Klassifizierungssyssteme (Classification Systems Being Able to Learn).* Berlin, Germany: Akademie-Verlag.

Zeller, C., & Schmid, U. (accepted). Rule learning from incremental presentation of training examples: Reanalysis of a categorization experiment. In *13th Biannual Conference of the German Cognitive Science Society (Bremen, Germany, Sept. 2016).*

Modelling Human Navigation: Cognitive Aspects of Obstacle Avoidance

Juan Purcalla Arrufi**

HCI chair, Universität Tübingen, Sand 14, 72076 Tübingen, Germany
juan.purcalla-arrufi@uni-tuebingen.de
http://www.hci.uni-tuebingen.de/

Abstract. Modelling human walking—specially obstacle avoidance—has applications on simulating human behaviour in emergency situations or implementing human acceptable navigation on robots. We research a model for human walking navigation, focused on the problem of avoiding another walking humans—the model should straightforward generalise to the avoidance of standing humans. We partition each avoidance trajectory in three stages: initiation of the trajectory adjustments, performance of the trajectory adjustments, finalisation of the trajectory adjustments. This division poses some questions that we want to tackle in this paper. First, what event triggers the trajectory adjustments, so that they are initiated at a certain moment (in time-space). Second, how the trajectory adjustments are performed: humans have two strategies to modify their trajectory—change the walking speed or direction— thus, we would like to know what makes choose one strategy over the other or to choose a combination of both.

1 Objective

We intend to provide a numerical model of the human avoidance of obstacles when walking. The model deals with two crossing humans: one of them is the *interferer*, i.e., he does not change at all the course or speed of his trajectory, the other is the *avoider*. As they approach—at constant velocity—the avoider performs adjusts his trajectory in order to avoid collision and reach the goal, which could have been reached in a straight line, were the interferer absent.

The numerical model should, on the one hand, reproduce the experimental results of the initiation of the trajectory adjustments—essentially that the smaller the crossing angle the latter the trajectory adjustment is initiated; and that the lower the speed the latter the trajectory adjustment is initiated. To that end we will use probabilistic inference based on a stochastic model for the interferer. On the other hand, the model should predict the way direction and speed adjustments are combined. A general result is that for obtuse angles pedestrians perform only direction adjustments; for acute–right angles pedestrians perform a speed adjustment in addition to the direction adjustment.

** Ph.D. supervisor Alexandra Kirsch; alexandra.kirsch@uni-tuebingen.de

2 Motivation

Understanding and modelling human navigation—pedestrian dynamics—is a task which began about 20 years ago to be an own field in science. This fact was prompted, among others, by two events: the founding of *Gait and Posture*—one of the leading journals in human locomotion—, and the seminal work of Helbing on pedestrian dynamics. Now we have an extensive research in crossing situations of pedestrians, which has persuaded us to deal with these situations.

A human navigation model may equip a robot moving in human environments with an acceptable navigation behaviour, what we also call *human-aware robot navigation*. Our ansatz is following: the most direct way to achieve that robots navigate in a human acceptable way is to make robots mimic human navigation. Human-aware robot navigation embraces numerous tasks of daily importance: assistive tasks in domestic environments, patrolling and surveillance, service delivery in health care institutions. Despite all research on robotics in the last four decades, human aware navigation has been just recently established as a discipline—this topic began to increasingly attract the attention of the scientific community in the year 2000. Consequently, many areas in this discipline still need both satisfactory solutions and a solid formalisation.

Apart of the benefits for robotics, modelling human navigation on crossing situations builds on cognitive science. Indeed, on the one hand, we argue that the adjustment of the avoidance trajectory is triggered by inference processes—the probability of colliding based on the inferred position distribution of the *interferer*. On the other hand, we research the causes of the avoidance strategy: what makes the *avoider* choose the direction or speed change in different proportions.

3 Method

We consider three consecutive parts in the trajectory of the *avoider*: initiation, performance, and finalisation of the trajectory adjustments. For both the initiation and the performance of the trajectory adjustments we test following methods as possible explanation the experimental results, and therefore as explanation for human behaviour.

3.1 Initialisation of Trajectory Adjustments

In a crossing situation we hypothesise that the trigger of the trajectory adjustments is the probability of the crossing distance being below a certain value (e.g., 0.5). The computations are based on the inferred probability distribution of the interferer at the crossing time, $t_\times$. We assume that the trajectory of the *interferer* is predicted as a stochastic Gaussian process whose expected value depends on the current interferer's velocity, i.e., $\langle \boldsymbol{x}(t) \rangle = \boldsymbol{v}_0 t$. By means of sequential Monte Carlo prediction we can infer the interferer's probability distribution at the crossing time $t_\times$ and, consequently, the probability of the crossing distance to be below the minimal crossing distance; which would trigger the initialisation of the trajectory adjustments.

3.2 Performance of Trajectory Adjustments

We assume that pedestrians perform trajectory adjustments based on three principles that we explain below: distances between humans, human speeds (maximal, minimal, typical), trajectory smoothness. When we require these three principles to determine the trajectory jointly, we expect to reproduce the observed combination of direction and speed adjustments.

Distances For static situations the most basic approach is the theory of *proxemics* that defines the acceptable distance intervals for the type of relation the static interferer has to the avoider (*public*, *social*, *personal*, or *intimate*).

When considering a moving interferer, i.e., kinematic situations, research abounds in crossing situations. One of the most remarkable results states that humans aim to keep a crossing distance (CD) of about 0.8 meters. Humans begin to adapt their trajectory based on the crossing distance they predict; their predictions assumes constant linear motion based on the current positions and velocities of both humans, k and l, $(\boldsymbol{x}_{k0},\ \boldsymbol{v}_{k0};\ \boldsymbol{x}_{l0},\ \boldsymbol{v}_{l0})$

$$\mathrm{CD}(\boldsymbol{x}_{k0}, \boldsymbol{v}_{k0}; \boldsymbol{x}_{l0}, \boldsymbol{v}_{l0}) = \min_{t \geq 0} \|(\boldsymbol{x}_{k0} + \boldsymbol{v}_{k0} t) - (\boldsymbol{x}_{l0} + \boldsymbol{v}_{l0} t)\| \tag{1}$$

Walking and Running Speeds The experiments of human locomotion have fixed the human values for walk speed: *slow* walking speed 1.15 m/s; *preferred* walking speed 1.41 m/s; *fast* walking speed 1.8 m/s; and *maximal* walking speed 2.3 m/s.

They have also found the limit for the transition into running modus, 2.05 m/s.

Smoothness Requirements Any trajectory is required to minimise the jerk's Root Mean Square (RMS) with certain boundary conditions in the interval $[t_1, t_2]$, e.g., $\boldsymbol{v}(t_1) = \mathbf{0}; \boldsymbol{x}(t_1) = \mathbf{0}$ and $\boldsymbol{v}(t_2) = \mathbf{0}; \boldsymbol{x}(t_2) = \boldsymbol{goal}$. Jerk minimisation is a general property of human motions: from arm displacements to walking trajectories.

We remark that this is a *global* (not local) requirement. It can only be fulfilled when the immediate future of the movement is, in some degree, predictable.

$$J = \langle j \rangle_2 = \left(\tfrac{1}{t_2 - t_1} \int_{t_1}^{t_2} \|\boldsymbol{j}(t)\|^2 \, dt \right)^{1/2} \quad \text{where } \boldsymbol{j}(t) = \dot{\boldsymbol{a}}(t) = \dddot{\boldsymbol{x}}(t) \tag{2}$$

Cognitive Complexity of Number Sequence Completion Problems: Evidence for Human Heuristics

Martin Hillebrand[1]

[1]Cognitive Systems Group, University of Bamberg, Germany
martin.hillebrand@stud.uni-bamberg.de

Abstract. Holzman, Pellegrino, and Glaser (1983) proposed a cognitive model for a number sequence completion problem solver. This model predicts that humans would not be able to solve a number sequence completion problem unless they have identified all rules. An empirical investigation was conducted where 41 participants were presented such problems with either one rule or two alternating rules. Results show that the difficulty is mostly governed by the single rule that is needed for extrapolating the next position. This indicates that the original model needs to be revised because humans are obviously able to skip goal-irrelevant rule identification activities.

1 Introduction

Number sequence completion problems (NSCPs) are a frequent component in intelligence tests, where they are supposed to measure inductive reasoning skills. Apart from psychometrics and the assessment of human intelligence, NSCPs are also used in Cognitive Modelling and in Artificial Intelligence nowadays, where different approaches are compared with regards to similarity to humans and with regards to maximum performance respectively. An overview over different computer models solving intelligence test problems can be found in Hernández-Orallo, Martínez-Plumed, Schmid & Siebers (2016). The research question of my Bachelor's thesis[1] is about different dimensions of cognitive complexity of NSCPs from a theoretical and an empirical point of view. In this submission, one of the empirical investigations conducted with human subjects will be presented.

A model that describes the affordances provided by a NSCP was developed by Holzman, Pellegrino, and Glaser (1983). It is a process model that is built upon the work of Simon and Kotovsky (1963) who were the first to describe a generic scheme for human reasoning on pattern detection. In their eyes, pattern detection consists of four hierarchically dependent phases. These are

1. Relations Detection: The subject needs to scan the sequence and form a hypothesis about a rule that describes the construction of the sequence.

[1] The complete content of my Bachelor's thesis can be retrieved from http://www.cogsys.wiai.uni-bamberg.de/theses/hillebrand/hillebrand.pdf

2. Discovery of Periodicity: The subject needs to determine the number of elements that build up a cycle pattern.
3. Completion of Pattern Description: The subject needs to complete the missing pattern within the periods from phase 2.
4. Extrapolation: The derived rule from phase 3 is applied to find out the solution.

Holzman et al. (1983) applied this scheme to the domain of solving NSCPs and developed a process model (Fig. 1). This model predicts that a human solving a NSCP will not proceed to the extrapolation of the solution until all rules (=relations) have been identified, but remain in a completion of pattern description loop. However, assuming that human information processing is based on the bounded rationality principle (Simon, 1956), it can be assumed that humans will skip the identification of rules which are irrelevant to the extrapolation step. Therefore, an empirical study that put the emphasize on the completion of pattern description process was conducted.

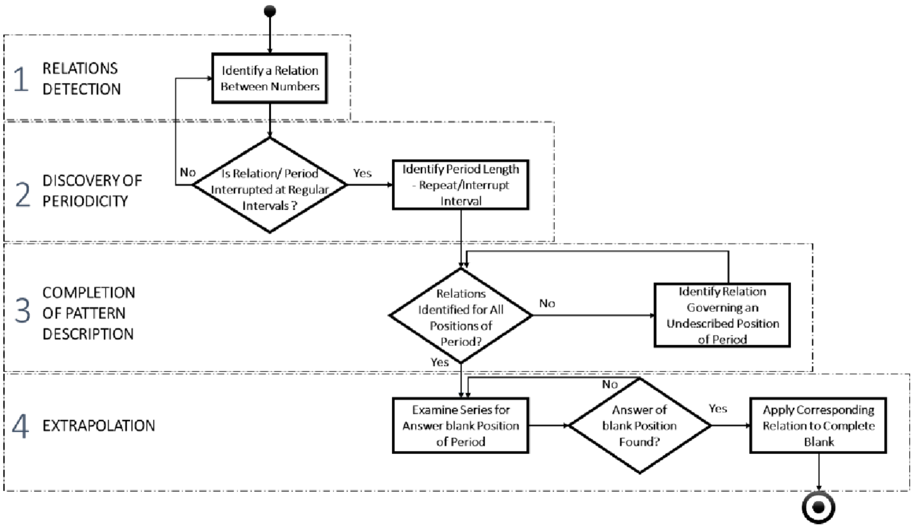


Fig. 1. The process model of a human solving a NSCP (adapted from Holzman et al., 1983).

2 Method

For this purpose, NSCPs were created that either can be described with one rule or with two different rules applied alternately. For every NSCP, six positions were given and the seventh needed to be extrapolated. For instance, the sequence 1, 4, 7, 10, 13, 16, … and 1, 2, 4, 8, 16, 32 … can be described in terms of one rule each, namely (+3) and (∗2) respectively, whereas the sequences 1, 4, 8, 11, 22, 25, … and 1, 2, 5, 10, 13, 26, … both need two rules (+3) and (∗2). These two sequences only differ in the order of the two rules and hence in the rule that is needed for the extrapolation step. According to the Holzman et al. process model, the two corresponding alternating sequences only differ with regards to the extrapolation step because it expects the human problem solver to identify all rules anyway. The specific rules that were used can be looked up in the complete text.

A total of N=41 subjects participated in this study. Every subject was assigned to one of eight different questionnaires each consisting of 39 NSCPs. Subjects were tested in groups in the university's computer rooms. NSCPs were presented individually and a time limit of three minutes was set for every single NSCP.

3 Results

In order to find evidence for our claim we looked at instances in our data set where a subject had to answer two single-rule-NSCPs and an alternating NSCP that included the rules from the two single-rule-NSCPs as rules. Then I looked at how different patterns of responses (correct/incorrect) for the two single-rule-NSCPs would go along with solving the alternating NSCP. The two single-rule-NSCPs were distinguished with regards to whether the rule was or was not the one that was needed for the extrapolation step of the alternating sequence. These results are presented in Table 1.

Table 1. Probability of a subject solving an alternating NSCP in dependence of the subject's pattern of answers of the two single-rule-NSCPs

	pattern of answers of the two single-rule-NSCPs of a person				
	relevant[a] rule	*irrelevant*[a] rule	number of instances of a pattern	number of instances with this person correctly solving the combined alternating NSCP	probability for this person correctly solving the combined alternating NSCP
1	correct	correct	495	310	62.626 %
2	correct	incorrect	109	73	66.972 %
3	incorrect	correct	108	17	15.741 %
4	incorrect	incorrect	26	3	11.538 %

[a] *relevant* or *irrelevant* refers to whether the rule of this single-rule-NSCP was relevant for the extrapolation of the solution in the alternating NCSP

In order to determine whether these differences in relative frequencies of correctly solved alternating NSCPs are significant, three Fisher's exact tests were calculated. Whereas no difference could be found for neither the comparison of $\text{correct}_{\text{relevant}}$/$\text{correct}_{\text{irrelevant}}$ with $\text{correct}_{\text{relevant}}$/$\text{incorrect}_{\text{irrelevant}}$ (row 1 with 2; p=.442) nor the comparison of $\text{incorrect}_{\text{relevant}}$/$\text{correct}_{\text{irrelevant}}$ with $\text{incorrect}_{\text{relevant}}$/$\text{incorrect}_{\text{irrelevant}}$ (row 3 with 4; p=.764), there was a significant difference between $\text{correct}_{\text{relevant}}$/$\text{incorrect}_{\text{irrelevant}}$ and $\text{incorrect}_{\text{relevant}}$/$\text{correct}_{\text{irrelevant}}$ (row 2 with 3; p<.001).

4 Discussion

It was assumed that, in contrast to what the Holzman et al. model predicts, humans can and do in fact skip the identification of rules that are irrelevant to the extrapolation step. When a person correctly solves a NSCP with a single rule, this performance is indicative of their success in both identifying this rule and applying this rule for extrapolation of the solution. Failure of solving such a single-rule-NSCP could be attributed to the failure of identifying or applying the rule. Assuming that an individual's ability to induce a specific rule from a number sequence is somewhat stable, one can use the performance of a single-rule-NSCP to capture an individual's ability to identify (and apply) a specific rule. Furthermore, this ability should help to detect the same rule in alternating NSCPs. When having estimated an individual's ability to induce two specific rules via two single-rule-NSCPs, one can use this to predict the likelihood to solve an alternating NSCP which contains both rules from the two single-rule-NSCPs.

In this study, results show that the probability of doing so is at its highest when the person solved the single-rule-NSCP whose rule was relevant for the extrapolation step (62.626% and 66.972%; rows 1 and 2 in Table 1) and it was lowest when this relevant single-rule-NSCP was not solved (15.741% and 11.538%; rows 3 and 4 in Table 1). In addition, no influence of solving the irrelevant single-rule-NSCP on performance of the alternating rule could be found whatsoever. In particular, it is interesting to observe that this is the case when knowing the relevant rule (row 1 and 2 in Table 1). The Holzman et al. model would state that people try to complete their description of a sequence before they extrapolate. This is apparently not the case since knowing the irrelevant rule is not indicative for humans to extrapolate the solution of an alternating NSCP.

Therefore, it can be concluded that it is rather the case that humans try to extrapolate the solution before they have identified all, particularly irrelevant rules and that they use efficient, goal-oriented heuristics in the context of NSCP. This suggests that the Holzman et al. model needs a revision. As demonstrated, humans do not determine all single rules in an alternating rule, but are able to proceed as soon as the relevant rule for the extrapolation has been identified. A modified process model should therefore include backward loops and fast-forward tracks.

Furthermore, inductive reasoning systems in the field Artificial Intelligence could profit from these results as skipping unnecessary rule detection increases effectiveness and efficiency, probably not only in the domain of NSCPs.

5 References

1. Hernández-Orallo, J., Martínez-Plumed, F., Schmid, U., Siebers, M., & Dowe, D. L. (2016). Computer models solving intelligence test problems: Progress and implications. *Artificial Intelligence*, *230*, 74-107.
2. Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. *Journal of Educational Psychology*, *75*(4), 603–618.
3. Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.
4. Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, *70*, 534–546.

Updating of Spatial Representations: Two pilot ERP Experiments

Efrosini (Frosso) Charalambous

Bartlett School of Architecture, UCL, 140 Hampstead Rd, London NW1 2BX,UK

frossoc.charalambous.12@ucl.ac.uk

Abstract. The present study investigates the moment of spatial reorientation as an instance of insightful problem solving in an attempt to answer the question: Is there a distinct pattern of brain activity reflecting the representational update in spatial problem solving? If we approach wayfinding as an instance of problem solving we can create intellectual bridges between lab-based experimental situations and real-world scenarios and adopt new perspectives on existing findings. Results from two pilot ERP experiments, one on orientation and one on puzzle solving, suggest a P3-like component might be the brain's signature of a representational update.

Keywords: Spatial problem solving, representational update, reorienation, EEG

1 Introduction

The contribution of the hippocampus in the reorientation of one's thinking, as observed in early studies on insights, has been suggested to be similar to that of reorientation in navigation[1]. Insightful solutions require restructuring of the initial misleading representation that had resulted in impasses. Mental impasses are also present when spatial updating is disrupted and our sense of direction is lost. Restructuring of the 'turned around' heading results in the Aha effect of reorientation. Waller&Hodgson's disorientation experiment provides evidence that disorientation causes a switch from the use of a temporary, egocentric representation of space to a more stable object-to-object representation [2]. The present study investigates brain activity during the representational update in two real-world scenarios: an orientation pointing task and a Sudoku puzzle. Sudoku puzzles involve sequential thinking and processing of spatial relations among elements. The aim was to identify moments of mental impasses: the subjective feeling of getting stuck and not knowing how to proceed. Operationalization of the moment of reorientation was based on Wang and Brockmole's experiment; when blindfolded subjects' heading was aligned with objects inside a room they were faster at pointing to indoor location than outdoor landmarks. But when aligned with outdoor locations they were equally fast [3]. These results suggest we operate on distinct representations over different timescales [4], e.g. in small and large-scale environments.

1.1 Brain Signatures of the Representational Update

ERP studies on insights have reported two main findings associated with the 'Aha!' moment: a negative deflection over the frontrocentral region, generated at the Anterior Cingulate Cortex (ACC) (detecting cognitive conflict and initializing the breaking of the mental set) [5, 6] and a positive wave at the parietoccipital electrodes generated at the parahippocampal gyrus (formation of novel associations) which according to several authors might be a P3-like component [6, 7]. Although the ACC meditates the successful restructuring of the representation, this signal was also present when subject *fail to comprehended* the provided answer [5]. A recent review on the insightful brain suggest that the effective switch of the problem's representation seems to depend on a visuospatial information-processing network [8]. Insight-like phenomena in visual perception (Necker cube, the Old/Young woman) elicited a similar parietal positivity (variant of the P3) at the moment of perceptual reversal. Researchers using manual responses as time reference, observed a positive component between -500ms -200ms before key press, which was interpreted as indicating conscious recognition of a perceptual switch [9]. According to Donchin's 'context updating theory', a restructuring of the mental model often occurs when there is a deviation from what is predicted or expected and it is conceivable that is manifested by the P3 [10]. Similar results are expected in the ERP pilot experiments.

2 Experiment

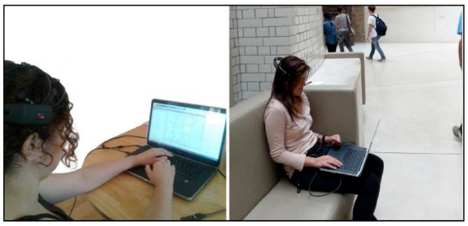


Fig. 1. Sudoku and Reorientation experiments

In the first experiment, four participants (familiar with Sudoku) were asked to solve three puzzles. After each puzzle they reported the times they got stuck before finding the next number. Subjects' self reports and reaction time (RT) were used for the distinction between test and control conditions. Subjects' experience of mental impasses resulted in greater reaction times, reflecting conflict in subjects' initial representation of possible solutions. These late correct responses were consider as the test Representational Update Condition and were compared to responses with RT<10sec. In the second experiment four participants (familiar with the area) followed a route in the V&A museum (fig.1). At four stop points subjects took a seat and were asked to rotate an image (showing a target location) using the arrow keys to align it with their heading orientation. In each stop point four images referred to target locations of the surrounding environment (e.g. South Kensington) and four to targets related with the indoor route. The outdoor-target correct response (deviation <30degrees) was hypothesize to require an update of the action-oriented egocentric representation with allocentric information and was considered test condition whereas correct route-related responses (locations updated with self-movement) as control.

2.1 Data Analysis

EEG signal was recorded with the Emotiv wireless system (14 electrodes). The EEG was down-sampled and band-pass filtered -0.16Hz -85Hz-by the device. The processing steps were carried out offline using EEGLAB and ERPLAB. Data were first preprocessed as suggested in the manual, using a high-pass filter at 0.2 Hz and a low pass filter at 40Hz to remove high frequency noise. An Independent component analysis ('runica' function) was applied to the continuous data. Stereotyped artifacts such as eye-blinks and eye-movement were detected and removed by the algorithm. Response-locked epochs, from -1000ms prior to keypress up to 200ms after were baseline corrected (whole epoch). Trials with deflections exceeding ±80 μV (peak-to-peak function) were marked and excluded from further analysis. Epochs of the different conditions (Aha/No-Aha) were averaged separately. A low-pass filter with a cut-off at 30 Hz was applied to the epoched data to remove further noise, following the recommendation of the toolbox's manual. Finally a grand-averaged waveform across subjects was produced for each bin and channel.

2.2 Results

The grand-averaged waveforms (fig.2) show that the test condition elicited greater positive deflections at the frontocentral, temporal and parietal areas. In the Sudoku task positive deflections were observed in the left frontocentral, temporal and parietal channels (FC5, T7, P7) around -600ms. The waveforms of the pointing task showed positive deflections in the time window -500ms to -200ms, greater at electrodes of the right hemisphere (FC6, T8, P8). Left parietal positivity for the problem-solving task was expected since patients with left parietal lesions have difficulties on visual constructive tasks. Whereas the difficulties of right parietal lobe patients in perceiving spatial relations among things and their tendency of getting lost justifies the greater right parietal positivity in the orientation task [11]. The parietal positivity probably signals the effective transformation of the representation; in Sudoku, reflecting the moment of restructuring the representation of elements' sequences after mental fixation and in the reorientation task updating of the current representation with spatial information of the outdoor environment, which does not occur automatically with self-movement.

3 Work in Progress

The research question that this PhD thesis is trying to answer is the following: If there is a distinct pattern of brain activity reflecting the Aha!moment of the representational update, then what are the visuospatial condition that may engender this ‘mental event ’? Since the pilot experiments suggest that the P3-like component might indeed reflect the restructuring of current mental representations, this signature can serve as an indicator of the Aha moment. Currently, I am designing a synchronised EEG/VR experiment in Unity3D where the evoked brain activity of test Vs control condition will be recorded and compare while subjects navigate 3D urban environments.

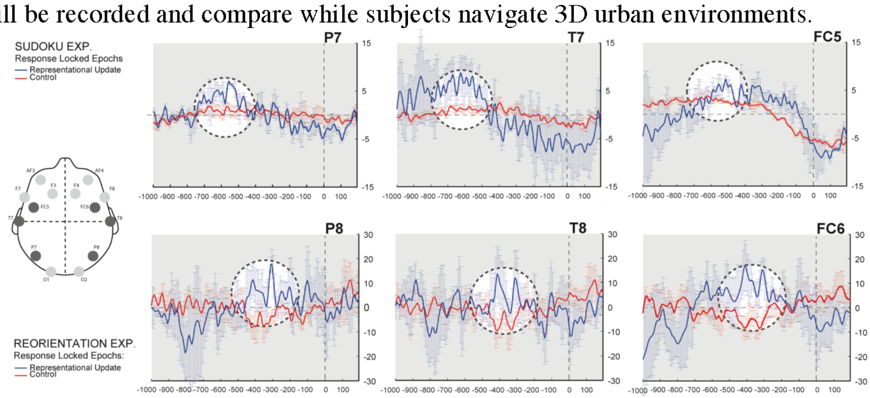


Fig. 2. ERP waveforms of Sudoku (first row) and Reorientation (second row) experiments.

Acknowledgments. This work has been partially supported by the Greek State Scholarship Foundation.

References

1. Luo J, Niki K. Function of hippocampus in "insight" of problem solving. Hippocampus. 13,316-23(2003)
2. Waller D, Hodgson E. Transient and Enduring Spatial Representations under Disorientation and Self-Rotation. Journal of Experimental Psychology: Learning, Memory, and Cognition. 32(4),867-82(2006)
3. Wang RF, Brockmole JR. Simultaneous spatial updating in nested environments. Psychon Bull Rev. 10(4),981-6(2003)
4. Burgess N. Spatial memory: how egocentric and allocentric combine. Trends in Cognitive Sciences. 10(12),551-7 (2006;)
5. Qiu J, Li H, Luo Y, Chen A, Zhang F, Zhang J, et al. Brain mechanism of cognitive conflict in a guessing Chinese logogriph task. NeuroReport. 17(6),679-82 (2006)
6. Wang T, Zhang Q, Li H, Qiu J, Tu S, Yu C. The time course of Chinese riddles solving: Evidence from an ERP study. Behavioural Brain Research. 199,278-82 (2009)
7. Qiu J, Zhang Q. “Aha!” effects in a guessing Chinese logogriph task: An event-related potential study. Chin Sci Bull. 53(3), 384-91(2008)
8. Shen W, Luo J, Liu C, Yuan Y. New advances in the neural correlates of insight: A decade in review of the insightful brain. Chinese Science Bulletin. 58(13),1497-511(2013)
9. Kornmeier J, Bach M. Ambiguous figures – What happens in the brain when perception changes but not the stimulus. Frontiers in Human Neuroscience. 6 (2012)
10. Donchin E. Surprise!… Surprise? Psychophysiology. 18(5),493-513(1981)
11. Dudchenko PA. Why People Get Lost. Oxford University Press, Oxford (2010)

The comprehension of verbal jokes: A visual-world study

Laura Israel & Evelyn C. Ferstl

IIG / Cognitive Science
Universität Freiburg

Abstract. Several studies have reported facilitation-effects for the comprehension of verbal jokes. To further investigate these findings, the following study used a Visual-World-Paradigm to analyze the time-course of joke comprehension compared to other ambiguous but non-funny stories. There were no significant differences in the temporal fixation-patterns between the two text categories, but joke-punchlines elicited a lower switching probability between the pictures. The results suggest that the facilitation effect in jokes can be attributed to a higher level of certainty about the joke interpretation caused by an affective feedback.

Keywords: Joke Comprehension, Verbal Humor, Visual-World-Paradigm.

1 Introduction

Most verbal jokes are built around a misunderstanding. Incongruity theories of joke comprehension hold that funny texts consist of a context that shapes the readers' expectations about the topic and about a likely continuation of the text. In the subsequent punchline (see Fig. 1), this expectation is disconfirmed and the situation model of the text needs to be updated.

To test this model, a number of studies analyzing reading times for verbal jokes have been conducted. For example, we reported facilitation for verbal jokes using eye movement monitoring during reading (Ferstl, Israel, & Putzar, 2016). We argued that the humorous content of jokes provides metacognitive feedback about the comprehension success, influencing rereading strategies more than immediate comprehension processes.

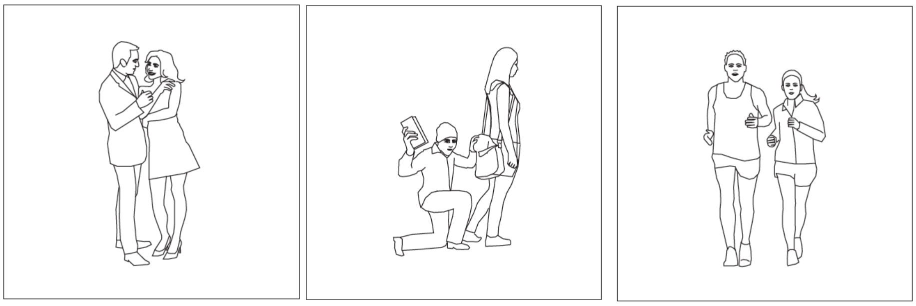
The following master thesis is trying to validate this feedback hypothesis and also wants to investigate how the affective component could facilitate the processing of jokes. To evaluate the content of comprehension, rather than comprehension difficulty, the present study employed a visual world paradigm (Huettig, Rommers, & Meyer, 2011; Kowatch, Whalen & Pexman, 2013), in which alternative situation models are represented pictorially. This method can give an insight into the temporal processing patterns of the two differential situation models. To disentangle the affective component of joke comprehension from cognitive revision processes, stories were included which required the updating of a situation model without being funny (Revision Stories, cf. Hunger, Siebörger & Ferstl, 2009).

Herbie to his pal Marc:

„Why have all these women stopped chasing after you lately?“ (joke)

„Great that you've managed to stay within the law lately.“ (revision)

Marc: „Well, life as a pickpocket turned out to be too risky in the long run.“



Competitor | Target | Distractor

Figure 1. Example stimuli for the joke condition and its control text. The text materials are translated from the original German.

For both experimental texts, we expected the viewing patterns during the context and punchline presentations to be reversed: while longer viewing times for the competitor were expected during the context presentation, the target was expected to be viewed more during the punchline presentation. Furthermore, we expected a replication of the joke facilitation effect: it was hypothesized that a facilitation would either show itself in an earlier shift to the target picture or in a lower probability of switching between the pictures in jokes compared to revision stories.

2 Methods

16 jokes and 16 revision stories were selected from Siebörger et al. (2006). These experimental stories and their coherent control stories were recorded using an automatic speech production system to minimize prosodic cues. For each text item three black and white line drawings were created (see Fig. 1). The visual stimuli consisted of a target-picture, representing the intended meaning of the story, an unrelated distractor, and one competitor that reflected the salient interpretation suggested by the context sentence for the revision and joke conditions. The same pictures were used for the experimental texts and their control texts. The position of the pictures on the screen (arranged in a triangle around a fixation cross) and the order

of the conditions were counterbalanced across conditions and participants.

26 native German speakers (13 male, 13 female) first viewed the pictures for 6 seconds, and then listened to the stories presented over headphones. After the end of the presentation, the pictures remained on the screen for 4 seconds, after which participants were prompted to perform a picture matching task. Using a mouse click, they chose the picture that best represented the meaning of the story. Finally, participants rated the funniness of the story on a 9-point-scale.

Eye movements were collected with an Eyelink-1000-System. Each of the three pictures was defined as a separate Interest Area (IA) for the analysis of the eye-tracking data. Four interest periods (IP) were defined: 6 seconds before story presentation (Start), the durations of the context (Context) and the punchline sentences (Punchline), resp., and 4 seconds before the prompt (End). The dependent variable was the percentage of viewing time on each IA in each IP.

3 Results

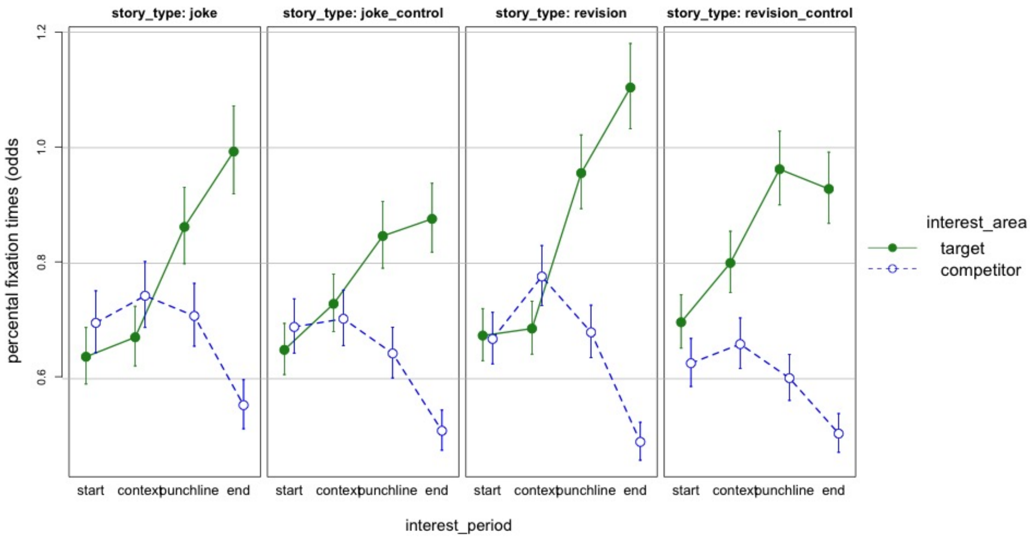


Figure 2. Percentage of the fixation times during the course of the trial for target and competitor picture.

The responses in the picture matching tasks were quite accurate (91%). Most errors were made in the joke condition (26%, range 41% - 100%). The funniness ratings confirmed the experimental manipulation, with jokes considered funnier (M=4.4) than revision and control stories (M=1.7).

The analysis of the eye tracking data (see Fig 2) confirmed that the three pictures were equally salient before the onset of the context sentence. For the experimental stories (revision and joke) the competitor was viewed more during the context sentence. As expected, the fixation durations on the target increased with the presentation of the punchline, and they were maximal before the performance of the matching task (End). There were no indications for an earlier switch to the target picture in jokes compared to revision stories. The analysis of the switching rate showed a trend towards a lower probability of switching between target and competitor during the punchline sentences in jokes compared to revision stories ($\chi^2(1) = 3.58$, $p = .058$).

4 Conclusions

In a previous reading study we reported a facilitation for jokes over non-funny text categories and argued that this effect might me due to a metacognitive feedback triggered by the perception of humor. The present study shows that jokes and the non-funny revisions stories show similar patterns in the temporal processing of the alternative situation models. These results suggest that the facilitation effect might not be caused by an earlier incongruity resolution in jokes compared to other ambiguous texts. Instead it could be argued, that the immediate feedback in jokes leads to a higher level of certainty about the correctness of the interpretation, as the punchlines of jokes elicited a lower switching probability in the current study. This interpretation would be in line with the findings of our previous reading-study, in which subjects reread non-funny texts more often than jokes. More fine-grained analyses will be presented to shed light on the time course of viewing within the interest periods considered. Most importantly, the fixations immediately following the disambiguating information in the punchline will be analysed (e.g., the word "pickpocket" in Fig. 1).

References

1. Ferstl, E. C., Israel, L., & Putzar, L. (in press / 2016). Humor facilitates text comprehension: Evidence from eye movements. *Discourse Processes*. Advance on-line publication: doi:10.1080/0163853X.2015.1131583.
2. Hunger, B., Siebörger, F., & Ferstl, E. C. (2009) Schluss mit lustig: Wie Hirnläsionen das Humorverständnis beeinträchtigen [Done with fun: How brain lesions affect humour comprehension]. *Neurolinguistik*, 21, 35-59.
3. Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137, 151-171.
4. Kowatch, K., Whalen, J. M., Pexman, P. M. (2013). Irony comprehension in action: A new test of processing for verbal irony. *Discourse Processes*, 50, 301-315. doi:10.1080/0163853X.2013.799934

Cognitive Sciences Strategies for Futures Studies (Foresight)

Ahmad Mahdeyan[1], Ebrahim SoltaniNasab[1]

[1] Both Ph.D. Student of Futures Studies (Foresight) at Strategic Management School, Supreme National Defense University, Tehran, IRAN
mahdeyan@iraneservice.com

Abstract. Developing the conceptual model of the origin of the idea of future scenarios leads to explore Cognitive Sciences (CS) strategies for Futures Studies (FS). This research will try to answer how scenario planning would benefit from CS by reshaping mental models? In other hand, how these explored strategies could develop the future oriented intelligence's machine? This is a vast amount of work to be considered. Modeling via abduction, chance-seeking via intervention on tacit knowledge, Acquiring useful information via causality grouping, Intelligence increase over time and idea blending are just the first examples, so we have a long way to go.

Keywords: Cognitive Science; Futures Studies; Mind Structure.

1 Introduction

Anticipation is increasingly at the heart of urgent contemporary debates, from climate change to economic crisis, bringing researchers together from across disciplines. The ability to anticipate in complex environments may improve the resilience of societies under threat from a global proliferation of agents and forces by articulating insecurities through anticipatory processes (1).

Popper (2) shows that three qualitative methods include Literature Review, Expert Panel and Scenarios are dominant methodologies tools to do foresight. Scenario thinking uses longer and broader views of possible futures to more clearly appreciate a world clouded by information overload, rapid change and uncertainty. Human's epistemological basis shows natural scenario building ability to tell stories about human life in the future (3) and a cognitive link to the time-oriented structure of the brain for perception. Our decisions about the future depend on how we think the world works. Scenarios are based on intuition, but crafted as analytical structures. We use Scenario planning artful via learning process to overcome barriers of creative thinking (4) via changing mental models for decision making. Scenarios are just different ideas about the future. We use them to guide us in exploring the future, widening perspectives (5), confronting assumptions, reshaping mental maps, etc.

Recently authors developed (6) the conceptual model of the origin of the idea of future scenarios by studying; idea ontology, the origin of creative thinking, idea nurturing in organizations, shaping the future time, scenario planning and idea's social network (global brain). Thus, on the basis of mind structure's role in causality blending of environmental and innate ideas based on copy principal and personal intelligence tools which attempts to use the benefits of the global brain via network collaboration, the new question emerged, hence, the research question that this proposal is trying to answer is the following;

How scenario planning would benefit from CS by reshaping mental models? In broader area, what will be CS strategies for FS? In other hand, how these explored strategies could develop the future oriented intelligence machine?

2 Methodology

This paper is a fundamental research type that makes theories for an applied science. Its analysis approach has been based on intuition-rational philosophy to explore new area of an interdisciplinary science by descriptive manner.

3 Literature Discussion

3.1 Cognitive Science (CS)

The scope of the study of cognition is broad; perception, mental representations, learning, mechanisms of reasoning, problem solving, intelligence and social psychology. CS emerged in the aftermath of World War II, driven by the invention of what we now recognize as information processing technologies (7). Today it also draws from fields such as philosophy, neuroscience, and anthropology, economics, epistemology, and the social sciences generally.

3.2 Futures Studies (FS)

Futures Studies or Strategic Foresight as its synonym is an interdisciplinary new generation of scientific attitude toward future that includes continuum from physical and biological issues to social and humanities subjects with most focus on technology's future and social changes in order to make desirable and sustainable future. Although some assume that FS is just a methodology, but based on more than thousands projects that successfully have been done in this field to represent a better understanding of future situations and taking appropriate actions in present (e.g., RAND and Shell (8)), we can demonstrate FS in 6 dimensions include; Presuppositions of the future time, Goals and objectives, Methodology, Outputs' materials, Futurists characteristics and Foresight horizon.

4 CS Strategies for FS

This is a vast amount of work to be considered and we have only sighted the first examples from the extended draft of this paper (9), and so have a long way to go.

4.1 Modeling could help Foresight to be Affordance via Abduction
  4.1.1 Modeling should not be very complicated although optimum level of stress is required for acquire attention in cognitive processes.
  4.1.2 Mathematical language could reduce models ambiguity
  4.1.3 FCM as a sample structure of modeling that could benefit FS
4.2 Chance-Seeking could benefits Foresight to find Wildcards and Weak Signals via Intervention on Tacit Knowledge to maximize Abducibility
4.3 Acquiring Useful Information via the Grouping based on Causality
4.4 Intelligence increase over time indicate improved Foresight by spending more time and use of collective intelligence as a Global Brain
4.5 Idea Blending and Interdisciplinary view could facilitate intuition of the Future Scenarios

5 Conclusion

Ideas and subjects that explained here are just introduction to this research main goal. We hope by wider and deeper investigations of CS Strategies for FS, the efficiency and the effectiveness of the present efforts to make better futures could be more productive. However some explored strategies of CS for FS could help us to develop machines that are able to planning for future and doing foresight.

Modeling could help foresight to be affordance via abduction but it should not be very complicated although optimum level of stress is required for acquire attention in cognitive processes. These developed models of FS could use as algorithms in artificial intelligence machines by facilitation of mathematical language. Using big data and also social network's recorded data by chance-seeking could benefits foresight in finding wildcards and weak signals via intervention on tacit knowledge to maximize abducibility and acquiring useful information via the grouping based on causality that today by ontological search engines is more easier. Then by collective intelligence as a global brain and blending ideas future scenarios could be narrated.

Thus these artificial intelligence machines could do actions for future in the present. This future oriented machine could use big data mining and global brain to get actions in the present in order to make better futures or prevent hazards of future uncertainties. So we can ask them to show us what is required to do now for future. Gaming software and Econometrics programs can be a simple example of present potential to achieve this future oriented intelligence machine as a science fiction idea.

6 References

1. **Poli, Roberto .** The Three Levels of Futures Studies: Forecast, Foresight and Anticipation. *UNESCO Chair in Anticipatory Systems, University of Trento.* 2015. roberto.poli@unitn.it.
2. **Popper, Rafael.** *Foresight Methodology.* Bonn, Germany : The Univercity of Manchester, 2008.
3. **Rhisiart, Martin .** *Scenario Building.* s.l. : Cardiff Business School, 2006.
4. **Chermack , Thomas J.** *Scenario Planning in Organizations: How to Create, Use, and Assess Scenarios.* s.l. : Berrett-Koehler Organizational Performance, 2011.
5. **Bentham, Jeremy B.** *Scenarios: An Explorer's Guide.* s.l. : Shell International BV, 2008.
6. *Origin Ontology of Future Scenario's Idea.* **Mahdeyan, Ahmad and SoltaniNasab, Ebrahim.** Tirupati : Sri Venkateswara University, 2016. Multilateral Cooperation: Emerging Global Scenario.
7. *Introduction to Cognitive Science.* **Ohlsson, Stellan.** 2015, Syllabus for CS491/PSCH 494.
8. **Borch, Kristian, Dingli, Sandra M and Jørgensen, Michael S.** *Participation and Interaction in Foresight: Dialogue, Dissemination and Visions.* s.l. : Edward Elgar Publishing, 2013. p. 352.
9. **Mahdeyan, Ahmad and SoltaniNasab, Ebrahim.** Cognitive Sciences Strategies for Futures Studies (Foresight). 2016. mahdeyan@iraneservice.com.

7 Biographical Sketch

Ahmad Mahdeyan is a junior scholar of Futures Studies (Strategic Foresight), which is an interdisciplinary new generation of scientific attitude toward future that includes continuum from physical and biological issues to social and humanities subjects with most focus on technology's future and social changes in order to make desirable and sustainable future. He has over ten years' experience in teaching business, management, economy and marketing and nearly two decades of work experience as customs expert. During these years he mentored and coached people to establish their own businesses and work with different cultures, he himself was also CEO at an online shopping center. Recently, he designed the Future Time Creation model during his Ph.D. theses and is planning to develop the conceptual model of the origin of the idea of future scenarios. Based on his experience, training is successful only when it leads to the behavioral changes. Our decisions about the future depend on how we think the world works. Opportunities are fleeting like clouds, so we should discover and seize the good ones. The idea of yesterday is the vision of today and the reality of tomorrow. So learn from the past, live for the present, and work for the future.

Index of Authors

——/ **A** /——

Abdel-Fatah, Ahmed M. H. . . . 131
Adam, Marion . . . 219
Agarwala, Aditya . . . 223
Albrecht, Rebecca . . . 201
Andresen, Erik . . . 237
Angerer, Benjamin . . . 173
Arrufi, Juan Purcalla . . . 245
Aumeistere, Anette . . . 187

——/ **B** /——

Bülthoff, Heinrich H. . . . 113
Baess, Pamela . . . 175
Baumann, Matthias-Philipp . . . 181
Beierle, Christoph . . . 99
Belardinelli, Anna . . . 121
Bergmann, Kirsten . . . 19, 131
Bergmann, Lasse T. . . . 183
Bermeitinger, Christina . . . 175
Bernardin, Keni . . . 219
Bläsing, Bettina E. . . . 71
Bleichner, Marin G. . . . 185
Bleichner, Martin G. . . . 197, 221
Blind, Felix . . . 183
Blumenschein, Stefanie . . . 121
Bordolo, Elias . . . 87
Bothe, Dietmar . . . 219
Bridgeman, Bruce . . . 77
Bry, Francois . . . 99
Burigo, Michele . . . 11, 105
Butz, Martin V. . . . 43, 121

——/ **C** /——

Carrera, Chiara . . . 187
Charalombous,Efrosini (Frosso) . . . 253
Cheema, Noshaba . . . 63
Chipofya, Malumbo . . . 131
Conroy, Ruth . . . 103
Czeszumski, Artur . . . 187

——/ **D** /——

Dablander, Fabian . . . 171
Dalton, Ruth Conroy . . . 1
Davies, Jim . . . 3
de Vladar, Harold P. . . . 143
Debener, Stefan . . . 185
Dehmel, Philipp . . . 189
Depner, Anamaria . . . 219
Dettling, Julian . . . 127
Dhesmuchade, Maja . . . 221
Dietsche, Lena . . . 15
Dimitrov, Todor . . . 219
Dittmer, Anke . . . 183
Doneit, Wolfgang . . . 219
Drimalla, Hanna . . . 191
Dshemuchadse, Maja . . . 213
Dziobek, Isabel . . . 191

——/ **E** /——

Engels, Marc Aurel . . . 219
Ernst, Marc . . . 119

——/ **F** /——

Faulhaber, Anja . . . 183
Fedor, Anna . . . 143
Ferstl, Evelyn C. . . . 15, 257
Fischer, Monika . . . 219
Fjaellingsdal, Tatiana Goregliad . . . 197
Foster, Celia . . . 113
Franke, Michael . . . 171
Franz, Ingo . . . 219
Freksa, Christian . . . 99, 145
Fridman, Lex . . . 63
Frisch, Simon . . . 199

——/ **G** /——

Güldenpenning, Iris . . . 59, 127
Gaerte, Philipp . . . 219
Gallagher, Shaun . . . 135
Gann, Mareike . . . 129, 195
Gehrig, Tobias . . . 219
Grage, Tobias . . . 199, 213
Greger, Karoline . . . 201
Gruebel, Jascha . . . 111
Grzeschik, Ramona . . . 103
Gunduz-Can, Rumeysa . . . 123

——/ **H** /——

Häussler, Jana . . . 163
Hölscher, Christoph . . . 31, 111, 229
Halbrügge, Marc . . . 51, 83, 169
Halfmann, Marc . . . 23, 215
Hardiess, Gregor . . . 215
Haschke, Robert . . . 119
Herzig, Joachim . . . 219
Hillebrand, Martin . . . 249
Himmelbach, Marc . . . 129, 195
Hofmann, Markus . . . 225
Holweger, Matthias . . . 209
Horeis, Chantal . . . 113
Hummel, Katrin . . . 213

——/ **I** /——

Indurkhya, Bipin . . . 147, 167
Innes, Anthea . . . 103
Ismail, Haythem O. . . . 131
Israel, Laura . . . 257

——/ **J** /——

Jacobs, Arthur M. . . . 225

Jacobsen, Thomas . . . 189, 225
Jaeger, Gerhard . . . 209
Jakubowitz, Tobias . . . 203
Jang, Juhee . . . 183
Jeczminska, Kinga . . . 95
Jost, Kerstin . . . 189
Juzek, Tom . . . 163

——/ K /——

König, Achim . . . 183
König, Peter . . . 67, 187
Kühnberger, Kai-Uwe . . . 131, 155
Kahl, Sebastian . . . 55
Kashkooli, Marjan Daneshvar . . . 205
Kern-Isberner, Gabriele . . . 99
Kiss, Nadina . . . 209
Kluth, Thomas . . . 11
Knauff, Markus . . . 99
Knoblic, Günther . . . 217
Knoblich, Günther . . . 47
Knoeferle, Pia . . . 11
Koester, Dirk . . . 59, 123
Kopp, Stefan . . . 19, 55
Kowollik, André . . . 203
Kraev, Simeon . . . 183
Krukar, Jakub . . . 115
Kruse, Andreas . . . 219
Kutz, Oliver . . . 131

——/ L /——

Lancier, Stephan . . . 27
Lanfenfeld, Vincent . . . 31
Lavelle, J. Suilin . . . 135
Lehner, Erich . . . 231
Linder, Stefan . . . 169
Llansola, Zoe Falomir . . . 131, 159, 193
Lohmann, Johannes . . . 43
Lohse, Jana . . . 219
Lotz, Alexander . . . 35
Luna-Rodriguez, Aquiles . . . 189, 225

——/ M /——

Mahdeyan, Ahmad . . . 261
Mallot, Hanspeter A. . . . 181, 215
Mallot, Haspeter A. . . . 23, 27
Maucher, Irene . . . 219
Meilinger, Tobias . . . 113
Meixner, Carmen . . . 183
Metzinger, Thomas . . . 5
Mikut, Ralf . . . 219
Mirkovic, Bojana . . . 185
Montecinos, Ernesto Andrés López . . . 187
Morisseau, Tiffany . . . 47
Moscatelli, Alessandro . . . 119
Murali, Supriya . . . 67

——/ N /——

Naceri, Abdeldjallil . . . 119
Newen, Albert . . . 139
Nosheen, Aalia . . . 183

——/ O /——

Oelliger, Michael . . . 143
Oliver, Eric . . . 193
Olteteanu, Ana-Maria . . . 151

——/ P /——

Patel-Groz, Pritty . . . 209
Pazzaglia, Francesca . . . 7
Perner, Josef . . . 139
Pipa, Gordon . . . 183
Prezensk, Sabine . . . 169
Prozmann, Viktoria . . . 23
Putze, Felix . . . 87, 219

——/ R /——

Räuker, Max . . . 183
Ragni, Marco . . . 99
Recanati, Francois . . . 139
Regier, Terry . . . 9
Renker, Johanna . . . 211
Ricken, Michael . . . 219
Rinkenauer, Gerhard . . . 211
Rosenholtz, Ruth . . . 63
Ruigendijk, Esther . . . 197
Russwinkel, Nele . . . 35
Rußwinkel, Nele . . . 169

——/ S /——

Sütfeld, Leon . . . 183
Santello, Marco . . . 119
Sauerbrei, Iris . . . 231
Schack, Thomas . . . 59, 123, 127
Scherbaum, Stefan . . . 197, 199, 213, 221
Schick, Wiebke . . . 181, 215
Schinazi, Victor . . . 111
Schlicht, Larissa . . . 183
Schlicht, Tobias . . . 135, 165
Schmid, Ute . . . 39, 203
Schmitz, Laura . . . 217
Schneider, Stefan . . . 131, 173
Schreiber, Cornell . . . 173
Schultheis, Holger . . . 11, 83, 109
Schultz, Tanja . . . 87
Schulz, Tanja . . . 219
Schulze, Timo . . . 219
Schwenke, Diana . . . 221
Schwering, Angela . . . 115
Sebanz, Natalie . . . 217
Seegelke, Christian . . . 127
Simon, Clarissa . . . 219
Singhal, Pulkit . . . 223
Smortchkova, Joulia . . . 139
SoltaniNasab, Ebrahim . . . 261
Sperber, Dan . . . 47
Sprengel, Michael . . . 225

Srivastava, Priyanka . . . 223
Stülpnagel, Rul von . . . 31, 201, 229
Stephan, Achim . . . 183
Stolzenburg, Frieder . . . 99
Szathmary, Eörs . . . 143
Szilagyi, Andras . . . 143

——/ **T** /——

Tamm, Sascha . . . 225
Thüring, Manfred . . . 227
Thrash, Tyler . . . 111
Timm, Silja . . . 183
Trapp, Anna Katharina . . . 227
Troja, Jörg . . . 231

——/ **V** /——

Vesper, Cordula . . . 47, 217, 221
Volpe, Massimo . . . 219

——/ **W** /——

Wächter, Max . . . 183
Wahn, Basil . . . 67, 187
Walker, Nina . . . 23
Walser, Moritz . . . 213
Watanabe, Katsumi . . . 113
Wendt, Mike . . . 189, 225
Wiener, Jan . . . 103
Wiese, Wanja . . . 91

——/ **X** /——

Xavier, Ann . . . 187

——/ **Z** /——

Zachar, Istvan . . . 143
Zeller, Christina . . . 39, 241
Zetzsche, Christoph . . . 63
Ziegler, Christof . . . 219