HEIDELBERG
EMPG 2019

50th Meeting of the European Mathematical Psychology Group

Heidelberg, August 5-7, 2019

Program and Abstracts

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Contents

Welcome

Dear Colleague,

It is our great pleasure to welcome you to the 50th Meeting of the European Mathematical Psychology Group held at Heidelberg University. We are pleased to present three excellent keynote speakers: Jeff Rouder, Carolin Strobl and Jennifer Trueblood. Furthermore, the conference program includes 45 talks and 17 posters. Because of a single-track program you won't miss any of them. New at this year's conference is the sports program. We invite you to attend the guided 10-minute sport session that takes place once a day.

We hope that the EMPG meeting will be stimulating and productive and that you will enjoy the atmosphere of the beautiful city of Heidelberg. Heidelberg University is the oldest university in Germany (founded in 1386). However, the conference will take place in one of the university's most modern buildings.

We want to thank the German Research Foundation (DFG, grant LE 4379/1-1) for their generous financial support. We further thank Stefan Radev, Mischa v. Krause and Edith v. Wenserski for their great support in organizing this conference. Finally, we would like to thank our conference assistants who will try to answer all your questions (look out for the red shirts).

Enjoy the conference!

The organizing committee

Veronika Lerche and Andreas Voss

General Information

Conference Organization

Organizers

- Veronika Lerche
- Andreas Voss

Co-Organizers

- Mischa v. Krause
- Stefan Radev
- Edith v. Wenserski

Assistants

- Felicitas Baumann
- Enrique Cifuentes
- Sarah Hladik
- Farina Lingstädt
- Yannick Roos
- Maximilian Theisen

Conference website and email

- https://www.psychologie.uni-heidelberg.de/empg2019/
- empg2019@psychologie.uni-heidelberg.de

Venue

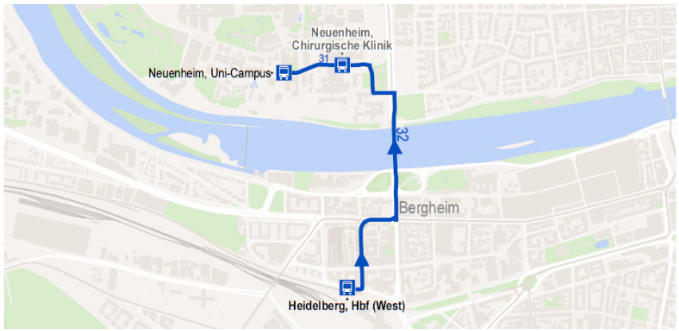
The conference takes place in the Marsilius-Kolleg building in Heidelberg, Germany (address: Im Neuenheimer Feld 130.1).
All talk sessions and keynotes will take place in the lecture hall ("Hörsaal"), located at the ground floor of the building. The poster session and the coffee breaks will be in the foyer, also located at the ground floor. The Welcome Reception on Sunday takes places in the "Clubraum", located at the first floor.

Travel

Heidelberg can easily be reached by train. Get your ticket to "Heidelberg Hbf" which is the main station of Heidelberg. By plane, you can reach Heidelberg via Frankfurt airport. There are bus and train connections from Frankfurt airport to Heidelberg.

From the train station the conference venue can be reached via a 20-minute walk, or an eight minute drive with the bus 32 to "Neuenheim Uni-Campus". You need to buy a ticket (2 zones, 2.60€) at the ticket machine before boarding the bus.

Walk from the central railway station to the bus stop "Heidelberg, Hauptbahnhof (West)" where the bus 32 departs from.



Route of bus 32 (direction: Neuenheim, Kopfklinik). You have to get off at "Neuenheim Uni-Campus" (please note that the bus changes its number to 31 at the stopover "Chirurgische Klinik").

Timetable for bus from Heidelberg Hbf to the conference venue

FAHRPLAN

VRN VERKEHRSVERBUND RHEIN-NECKAR

Handy: http://mobil.vrn.de Internet: http://www.vrn.de

http://m.vrn.de/1214

32

Direction:
Neuenheim, Kopfklinik

Heidelberg, Universitätsplatz
Heidelberg, Hbf (West)
- Betriebshof
Neuenheim, Jahnstraße
- Chirurgische Klinik
- Uni-Campus
- Botanischer Garten
- Zoo/Medizinische Klinik
- Jugendherberge
- Studentenwohnheim
- Kopfklinik

02 05 06 08 09 11 12 13 14

Time	Monday – Friday	Time	Saturday	Time	Sun- and Holiday
5	36[D] 46[D] 56[D]	5	45[D]	5	
6	08[D] 16[D] 28[D] 36[D] 48[D] 56[D]	6	15[D] 45[D]	6	45[D]
7	08[D] 16[D] 29[D] 39[D] 49[B D] 49[C D] 59[D]	7	15[D] 45[D]	7	15[D] 45[D]
8	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	8	15[D] 45[D]	8	15[D] 45[D]
9	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	9	06[D] 26[D] 46[D]	9	15[D] 45[D]
10	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	10	06[D] 26[D] 46[D]	10	15[D] 45[D]
11	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	11	06[D] 26[D] 46[D]	11	15[D] 45[D]
12	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	12	06[D] 26[D] 46[D]	12	15[D] 46[D]
13	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	13	06[D] 26[D] 46[D]	13	06[D] 26[D] 46[D]
14	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	14	06[D] 26[D] 46[D]	14	06[D] 26[D] 46[D]
15	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	15	06[D] 26[D] 46[D]	15	06[D] 26[D] 46[D]
16	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	16	06[D] 26[D] 46[D]	16	06[D] 26[D] 46[D]
17	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	17	06[D] 26[D] 46[D]	17	06[D] 26[D] 46[D]
18	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	18	06[D] 26[D] 46[D]	18	06[D] 26[D] 46[D]
19	09[D] 19[D] 29[D] 39[D] 49[D] 59[D]	19	06[D] 26[D] 46[D]	19	06[D] 26[D] 46[D]
20	09[D] 19[D] 31[A] 41[A] 45[D]	20	06[D] 26[D] 45[D]	20	06[D] 26[D] 45[D]
21	15[D] 45[D]	21	15[D] 45[D]	21	15[D] 45[D]
22	15[D] 45[D]	22	15[D] 45[D]	22	15[D] 45[D]
23	15[D] 45[D]	23	15[D] 45[D]	23	15[D] 45[D]
0	15[D]	0	15[D]	0	15[D]

A: Terminal stop at „Heidelberg, Betriebshof"
B: On school holidays only
C: On school days only
D: Continuing as number 31 at stop „Chirurgische Klinik"

Further and current information: www.rnv-online.de

Rhein-Neckar-Verkehr GmbH, Möhlstr.27, 68165 Mannheim, www.rnv-online.de, 0621/465-4444, Tel.: 0621/465-4444

Timetable for bus from the conference venue to Heidelberg Hbf

FAHRPLAN

VRN VERKEHRSVERBUND RHEIN-NECKAR

Handy: http://mobil.vrn.de Internet: http://www.vrn.de

http://m.vrn.de/6144

Direction:
Universitätsplatz - Heidelberg, Bismarckplatz

32

Neuenheim, Kopfklinik
Neuenheim, Uni-Campus
- Chirurgische Klinik
- Jahnstraße
Heidelberg, Betriebshof
- Hauptbahnhof (West)
- Stadtwerke
- Stadtbücherei
- Seegarten
- Bismarckplatz
- St. Vincentius-Krankenhaus
- Kongresshaus
- Marstallstraße
- Universitätsplatz

01 02 05 08 11 13 14 16 21 23 24 26

Time	Monday – Friday	Time	Saturday	Time	Sun- and Holiday
5	48	5	51	5	
6	08 28 46	6	21 51	6	
7	08 28 48 58	7	21 51	7	21 51
8	08 18 28 38 48 58	8	21 48[C]	8	21 51
9	08 18 28 38 48 58	9	08[C] 28[C] 48[C]	9	21 51
10	08 18 28 38 48 58	10	08[C] 28[C] 48[C]	10	21 51
11	08 18 28 38 48 58	11	08[C] 28[C] 48[C]	11	21 51
12	08 18 28 38 48 58	12	08[C] 28[C] 48[C]	12	21[C] 48[C]
13	08 18 28 38 48 58	13	08[C] 28[C] 48[C]	13	08[C] 28[C] 48[C]
14	08 18 28 38 48 58	14	08[C] 28[C] 48[C]	14	08[C] 28[C] 48[C]
15	08 18 28 38 48 58	15	08[C] 28[C] 48[C]	15	08[C] 28[C] 48[C]
16	08 18 28 38 48 58	16	08[C] 28[C] 48[C]	16	08[C] 28[C] 48[C]
17	08 18 28 38 48 58	17	08[C] 28[C] 48[C]	17	08[C] 28[C] 48[C]
18	08 18 28 38 48 58	18	08[C] 28[C] 48[C]	18	08[C] 28[C] 48[C]
19	08 18 28 38 48 58	19	08[C] 28[C] 48[C]	19	08[C] 28[C] 48[C]
20	08[A] 18[A] 25 38[A] 51	20	08[A] 25 51	20	08[A] 25 51
21	21 51	21	21 51	21	21 51
22	21 51	22	21 51	22	21 51
23	21 51	23	21 51	23	21 51
0	21 55[B]	0	21 55[B]	0	21 55[B]

A: Terminal stop at „Heidelberg, Betriebshof“ **B:** Terminal stop at „Heidelberg, Hbf (West)
C: Terminal stop at „Heidelberg, Universitätsplatz“

Further and current information: www.rnv-online.de

Rhein-Neckar-Verkehr GmbH, Möhlstr.27, 68165 Mannheim, www.rnv-online.de, 0621/465-4444, Tel.: 0621/465-4444

Accomodation

There are a lot of hotels close to the conference venue, both in the city center and around the train station.

Registration Desk

At the Registration Desk, you can register to the conference, pick up your conference materials and ask for any assistance. The Registration Desk can be found in the foyer of the Marsilius-Kolleg building. Opening hours during the conference (Monday to Wednesday): 08:00 am to 06:00 pm. Please note that you can also obtain your conference materials in advance during the Welcome Reception on Sunday evening.

Welcome Reception

The Welcome Reception will take place in the “Clubraum“, located at the first floor of the Marsilius-Kolleg building, on Sunday, August 4, from 06:00 to 08:00 pm. You will have the opportunity to meet other conference participants, pick up your conference materials and enjoy some refreshments and snacks.

Presentation Guidelines

Talks

Talks are scheduled for 20 minutes, including 15 minutes for presentation and 5 minutes for discussion. All talks will take place in the lecture hall (“Hörsaal“). You have the possibility to connect your own computer or tablet to the presentation equipment. If you plan to use the local computer (Windows 10), please hand your presentation in PPT or PDF format to the technical assistant in the lecture hall before the session starts.

Posters

The poster session will take place in the foyer on Monday from 04:30 to 05:30 pm. The poster boards are 200 cm high and 108 cm wide. Please attach your poster (A0) to the board before 01:00 pm. Posters

must be attached with adhesive strips, which will be provided by the conference assistants.

Internet Access

You can use *eduroam* with your university credentials and password. In case you cannot connect via *eduroam*, please ask the assistants at the Registration Desk for a guest login.

Sports Program

A new feature in this year's conference is the sports program. Conference days are long and exhausting and the sports units will hopefully help you to get some fresh energy for the afternoon sessions. On Monday to Wednesday at 02:50 pm a trainer from the sports department of Heidelberg University will conduct a ten minute sport session. The session will take place in the lecture room of the Marsilius-Kolleg. You do not need to bring sports clothing, your conference outfit is just fine.

Poster Session

Do not miss this year's poster session! In addition to getting to know interesting new research, you will have the opportunity to enjoy refreshments and snacks. Thereby, you can also get prepared for the Philosopher's Walk starting right after the poster session (see below).

Philosopher's Walk & Dinner

On Monday after the poster session—if the weather permits—we will take a walk along Heidelberg's famous *"Philosopher's Walk"*. We will meet at 05:40 pm in the foyer of the Marsilius-Kolleg. From there, we will first walk along the Neckar river and then proceed to the Philospher's Walk (altitude difference of about 100 meters) and finally cross the *Alte Brücke* (Old Bridge).

We hope that like professors and philosophers of the past, you will gain new research ideas inspired by walking this path with its stunning views upon the old part of the city and the castle. For further information about the Philosopher's Walk, see: https://www.heidelberg-marketing.de/en/experience/sights/philosophenweg.html.

The walk will take about 1 to 2 hours and end at about 07:30 pm at the restaurant *Palmbräu Gasse* (address: Hauptstraße 185) where we reserved tables to have dinner together (self-pay). For the non-philosophers or non-walkers: Feel free to join us at the restaurant.

Walk from the Conference Venue to the *Palmbräu Gasse*

Conference Dinner

Conference Dinner will take place on Tuesday starting at 08:00 pm in the *BräuStadl* (Berliner Straße 41), a typical German restaurant. The *BräuStadl* is just a ten minute walk from the conference building. Conference Dinner is included in the registration fees. Please remember to bring your dinner voucher.

Walk from the conference venue to the *BräuStadl*

Lunch and Coffee Breaks

During coffee breaks, coffee and snacks will be provided in the foyer of the Marsilius-Kolleg building. You can get free lunch at the Mensa (address: Im Neuenheimer Feld 304) using the vouchers that are included in the conference materials (one meal and one drink per day). The Mensa is a five minute walk from the Marsilius-Kolleg building.

Walk from the conference venue to the Mensa

Abstracts

Keynotes

Qualitative vs. Quantitative Individual Differences: Implications for Cognitive Control

Jeff Rouder

University of California, Irvine

Consider a task with a well-established effect such as the Stroop effect. In such tasks, there is often a canonical direction of the effect—responses to congruent items are faster than incongruent ones. And with this direction, there are three qualitatively different regions of performance: (a) a canonical effect, (b) no effect, or (c) an opposite or negative effect (for Stroop, responses to incongruent stimuli are faster than responses to congruent ones). Individual differences can be qualitative in that different people may truly occupy different regions; that is, some may have canonical effects while others may have the opposite effect. Or, alternatively, it may only be quantitative in that all people are truly in one region (all people have a true canonical effect). Which of these descriptions holds has two critical implications. The first is theoretical: Those tasks that admit qualitative differences may be more complex and subject to multiple processing pathways or strategies. Those tasks that do not admit qualitative differences may be explained more universally. The second is practical: it may be very difficult to document individual differences in a task or correlate individual differences across task if these tasks do not admit qualitative individual differences. In this talk, I develop trial-level hierarchical models of quantitative and qualitative individual differences and apply these models to cognitive control tasks. Not only is there no evidence for qualitative individual differences, the quantitative individual differences are so small that there is little hope of localizing correlations in true performance among these tasks.

A Statistician's Botanical Garden - The Ideas behind Decision Trees and Random Forests

Carolin Strobl

University of Zurich

Classification and regression trees, model-based trees and random forests are powerful statistical methods from the field of machine learning. They have been shown to achieve a high prediction accuracy, especially in big data applications with many predictor variables and complex association patterns (such as nonlinear and higher-order interaction effects). While individual trees are easy to interpret, random forests are "black box" prediction methods. They do, however, provide variable importance measures, that are being used to judge the relevance of the individual predictor variables. The aim of this presentation is to introduce the rationale behind trees, model-based trees and random forests, to illustrate their potential for high-dimensional data exploration in psychological research, but also to point out limitations and potential pitfalls in their practical application.

Urgency, Leakage, and the Relative Nature of Information Processing in Decision-making

Jennifer Trueblood

Vanderbilt University

Over the last decade, there has been a robust debate in decision neuroscience and psychology about what mechanism governs the time course of decision making. Historically, the most prominent hypothesis is that neural architectures accumulate information over time until some threshold is met, the so-called Evidence Accumulation hypothesis. However, most applications of this theory rely on simplifying assumptions, belying a number of potential complexities. Is stimulus information perceived and processed in an independent manner or is there a relative component to information processing? Does urgency play a role? What about evidence leakage? While the latter questions have been the subject of recent investigations, most studies to date have been piecemeal in nature, studying one aspect of the decision process or another. Here we develop a modeling framework, an extension of the Urgency Gating Model, in conjunction with a changing information experimental paradigm to jointly probe these aspects of the decision process. Using state-of-the-art Bayesian methods to perform parameter-based inference, we demonstrate 1) information processing is relative with early information influencing the perception of late information, 2) time varying urgency and evidence accumulation are of roughly equal importance in the decision process, and 3) leakage is present with a time scale of ~250ms. This is the first such study to utilize a changing information paradigm to jointly and quantitatively estimate the temporal dynamics of human decision-making.

Talks

Unique skills assessment via minimal competence models

Pasquale Anselmi[a] , Jürgen Heller[b] , Luca Stefanutti[a] & Egidio Robusto[a]

[a] University of Padua, [b] University of Tübingen

Assessing the latent competence state of an individual (i.e., the subset of available skills) from the responses given to a set of items requires to first infer the latent knowledge state (i.e., the subset of mastered items). Unique assessment need to be based on a one-to-one correspondence between the collection of all the competence states occurring in the population and the collection of the knowledge states delineated by those competence states via a competence model. The latter includes a mapping associating to each item one or more subsets of skills, each of which is sufficient for solving the item. The talk explores the conditions under which a competence model can be constructed that results in the aforementioned one-to-one correspondence. A procedure is proposed that, for a fixed collection of competence states, allows for constructing competence models that differ from one another with respect to the items and the subsets of skills assigned to them, but are the same with respect to the assessed competence states. The resulting competence models are minimal (i.e., no item can be deleted without altering the assessment). The construction of conjunctive (for each item, there is a unique set of skills) and disjunctive (for all the items, each of the skills assigned to an item is sufficient for solving it) competence models is considered. Various applications of the suggested procedure will be presented. They cover different scenarios, such as developing a test from scratch, and improving or shortening an existing test.

The Dynamics of Decision Making During Goal Pursuit

Timothy Ballard[a] , Andrew Neal[a] , Simon Farrell[b] & Andrew Heathcote[c]

[a] University of Queensland, [b] University of Western Australia, [c] University of Tasmania

Goal pursuit can be thought as a series of interdependent decisions made in an attempt to progress towards a performance target. Whilst much is known about the intra-decision dynamics of single, one-shot decisions, far less is known about how this process changes over time as people get closer to achieving their goal and/or as a deadline looms. For example, people may respond to a looming deadline by either increasing the amount of effort they apply or by changing strategy. We have developed an extended version of the linear ballistic accumulator model that accounts for the effects that the dynamics of goal pursuit exert on the decision process. In this talk, I describe recent studies from our lab that test this model. In each study, participants performed a random dot motion discrimination task in which they gained one point for correct responses and lost one point for incorrect responses. Their objective was to achieve a certain number of points within a certain timeframe (e.g., at least 30 points in 40 seconds). Preliminary results suggest that decision thresholds were highly sensitive to deadline, such that people prioritised speed over accuracy more strongly as the time remaining to achieve the goal decreased. The decision process was also sensitive to the amount of progress that remained before the goal was achieved, the difficulty of the decision, the incentive for goal achievement, and whether the goal was represented as an approach goal or an avoidance goal. These findings illustrate the sensitivity of decision making to the higher order goals of the individual, and provides an initial step towards a formal theory of how these higher level dynamics play out.

The disjunction effect in two-stage gamble experiments

Jan Broekaert[a] , Jerome Busemeyer[a] & Emmanuel Pothos[b]

[a] Indiana University, [b] City, University of London

In 1992, Tversky and Shafir showed that Savage's rational axiom of decision making under uncertainty, called the 'Sure Thing' principle, was empirically falsified in a two-stage gamble experiment. It revealed that subjects would take a second stage gamble for both possible outcomes of the first stage gamble, but would not do so when no information was available on the outcome of the first stage gamble. They called this 'violation of the Sure Thing principle' a 'Disjunction Effect', which they further identified by the 'only stop on unknown outcome condition' gamble pattern outnumbering the 'play on all outcome conditions' gamble pattern. Since each responder gamble pattern is the outcome of an individual stochastic decision process, we will interpret the Disjunction Effect as an group aggregate level violation of the Law of Total Probability. The violation of the Law of Total probability emerges from the probability distribution of the eight possible gamble patterns each with its proper tendency to inflate or deflate the marginal gamble probability under Unknown outcome condition.Subsequent research in the literature has reported difficulty replicating the Disjunction Effect. We replicated this experimental paradigm in an online study (N = 1119) in which we adapted the range of payoff amounts, and controlled the order of the two stage gambles with, or without, information on the outcome of the first stage gamble. We introduced an operational measure for risk aversion based on the total number of accepted single stage gambles, and an Inflation-Deflation score based on the responder's formed gamble patterns. Surprisingly, we found that (a) less risk averse responders produced no disjunction effect, but (b) more risk averse responders produced a violation of the Law of Total Probability, where the direction of this violation depended on the order of the uninformed and informed gambles. We developed three models, a logistic model, a Markov model and a quantum-like model for this gamble decision process which shared the same basic features i) the decision is a dynamic process driven by utility, ii) implementation of a contextual influence on the belief state within its period and flow order of the outcome conditions, iii) implementation of a carry-over influence on the belief state from first to second period and, 4) a re-initialization principle of the contextual belief

state in each period of the flow order. Partitioning of the responders into seven Inflation-Deflation score ranges (symmetric [-2,2], [-1,1], [0,0]; positive]-2,2], [0,1]; and negative [-2,2[, [-1,0] reach) shows specific subgroups of the more risk averse responders at the root of the violation the Law of Total Probability. A log-likelihood model performance comparison indicates these specific subgroups are typically best replicated by the quantum-like process model.

A Hidden Markov Model approach to model Mouse-Tracking Data

Marco D'Alessandro[a] , Luigi Lombardi[a] & Antonio Calcagni[b]

[a] University of Trento, [b] University of Padua

Computer mouse-tracking recording techniques can provide a behavioral measure of the cognitive dynamics involved in a wide range of cognitive processes such as, for example, decision-making, categorization, and language processing. The richness of the spatial and temporal data offered by mouse trajectories allow to test hypothesis regarding the cognitive mechanisms underlying the time-course of decisions and behavioral responses. In the present work, we propose a unified approach to analyse mouse trajectories within a dynamic latent state framework which accounts for both motor(-spatial) and mental processes information at the same time. In particular, our model relies upon a discrete representation of mouse trajectories events, as well as of the corresponding mental processes involved. Here, the main purpose is to map the evolution of observed mouse-tracking recordings with the evolution of a cognitive latent state process underlying these observations. The characteristics and potentials of our approach are illustrated using a case study on categorization task.

Extracting Partially Ordered Clusters from ordinal polytomous data: A comparison between *k*-modes and *k*-median algorithms

Debora de Chiusole, Andrea Spoto & Luca Stefanutti

University of Padua

In the framework of data mining, clustering is a well-known statistical technique that consists of grouping a collection of data points. Given a set of data points, a certain clustering algorithm can be used for classifying each point into a specific cluster. Among the different algorithms developed to this aim, there are *k*-mode, *k*-median and *k*-means. All of them consist of the accomplishment of the following two Steps: (1) Classify the data points into a certain number of different clusters *K*, by using a certain dissimilarity measure; (2) Adjust each cluster *K* so that the within mean discrepancy between *K* and its group of data is minimized. The difference among the three algorithms is the type of measure on which these two steps are performed. In knowledge space theory, knowledge states can be considered as partially ordered clusters of individuals that form a knowledge structure. This theory has been recently extended to cover the quite common situation of polytomous items. An adaptation of the *k*-median algorithm is proposed for extracting polytomous structures from the data. The proposed algorithm is an extension of *k*-modes to ordinal data in which the Hamming distance is replaced by the Manhattan distance in Step (1), and the central tendency measure is the median rather than the mode in Step (2). A series of simulation studies and an empirical application have been carried out for comparing the performances of the two algorithms. Results show that there are theoretical and practical reasons for preferring the *k*-median to the *k*-modes algorithm, whenever the responses to the items are measured on an ordinal scale. This is because the Manhattan distance is sensitive to the order on the levels, while the Hamming distance is not. The possibility of comparing the performances of these two algorithms with those of *k*-means concludes the discussion.

From Italian menus to resolutions of learning spaces

Jean-Paul Doignon

Université Libre de Bruxelles

A typical two-stage process governs the selection of meal items in an Italian menu. The process admits a direct formalization in terms of choice spaces (the latter are defined as in Plott, 1973). Path-independent choice spaces are cryptomorphic to learning spaces (Koshevoy, 1999). There results a construction of a new learning space for a learning space (the base) given together with a family of learning spaces (the fibers) indexed by the elements of the base; we call the outcome a resolution. The construction is similar to the classical composition of hypergraphs due to Chein, Habib & Maurer (1981), Möhring & Radermacher (1984), Ehrenfeucht & McConnell (1994). We investigate resolutions of learning spaces, reporting first results on indecomposable learning spaces. The talk is based on on-going, joint work with Domenico Cantone, Alfio Giarlotta and Stephen Watson.

A comparison of conflict diffusion models in the flanker task through pseudo-likelihood Bayes factors

Nathan Evans[a] & Mathieu Servant[b]

[a] University of Amsterdam, [b] Université de Franche-Comté

Conflict tasks are one of the most widely studied paradigms within cognitive psychology, where participants are required to respond based on relevant sources of information while ignoring conflicting irrelevant sources of information. The flanker task, in particular, has been the focus of considerable modeling efforts, with only three models being able to provide a complete account of empirical choice response time distributions: the dual-stage two-phase model (DSTP), the shrinking spotlight model (SSP), and the diffusion model for conflict tasks (DMC). Although these models are grounded in different theoretical frameworks, can provide diverging measures of cognitive control, and are quantitatively distinguishable, no previous study has compared all three of these models in their ability to account for empirical data. Here, we perform a comparison of the precise quantitative predictions of these models through Bayes

factors, using probability density approximation to generate a pseudo-likelihood estimate of the unknown probability density function, and thermodynamic integration via differential evolution to approximate the analytically intractable Bayes factors. We find that for every participant across three data sets from three separate research groups, DMC provides an inferior account of the data to DSTP and SSP, which has important theoretical implications regarding cognitive processes engaged in the flanker task, and practical implications for applying the models to flanker data. More generally, we argue that our combination of probability density approximation with marginal likelihood approximation provides a revolutionary step for the future of model comparison, where Bayes factors can be calculated between any models that can be simulated.

Fuzzy Item Ambiguity Analysis in psychological testing and Measurement

Hojjatollah Farahani & Parviz Azadfallah

Tarbiat Modares University, Tehran

Item Ambiguity is a significant factor in psychological testing and measurement. Item Ambiguity is an unavoidable part of psychological testing and assessment. This item statistic has received no attention in classical test theory (CTT) and item response theory (IRT) of psychometrics so far. All of the items of a psychological test are of the degree of ambiguity. This is able to influence on item discriminant coefficient, test validity, reliability, and diagnostic accuracy. It can be a part of the item fairness in the achievement tests as well. The ambiguity of an item is defined as the degree of the perceived fuzziness of the content of that item (Farahani, Wang & Oles, 2018). To determine the item ambiguity of a test, this paper recommended a 5-stage process using fuzzy logic theory. In this paper, this method was presented and illustrated the calculation steps with a numerical example.

Generalizing the Memory Measurement Model to n-AFC recognition retrievals

Gidon T. Frischkorn & Klaus Oberauer

University of Zurich

The memory measurement model (M3; Oberauer & Lewandowsky, 2018) assumes that different categories of representations in working memory get activated through distinct processes within working memory. Transforming the activation of the different item categories into their respective recall probabilities allows to estimate the contributions of different memory processes to working memory performance. In this talk, I will outline the specific formulation of the M3 and present a generalization of the model for n-alternative forced choice (n-AFC) recognition retrievals. In n-AFC recognition retrievals participants are forced to choose their responses from a set of given options, instead of freely recalling items from memory. This reduction of retrieval options is particularly useful when experimenters are interested in both the accuracy and response time of retrieval from working memory. In this, a generalization of the M3 for n-AFC retrievals opens up the possibility to use evidence accumulation models such as the linear ballistic accumulator model (Brown & Heathcote, 2008) or the diffusion model (Ratcliff, 1978) as choice rule. This would enable researchers to investigate the effects of different cognitive processes within working memory on both the quality of representations and their speed of memory retrieval. The results from simulations show that parameters from the M3 can be well recovered acceptably even for 2-AFC retrievals varying the presented lure across all possible categories of memory representations. In this, the M3 model provides an interesting approach to the measurement of theoretically founded parameters for different cognitive processes in working memory. Moreover with the generalization to n-AFC recognition retrievals the model can be applied in a wide range of different tasks and paradigms.

References:

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cognitive psychology, 57*, 153–178.

Oberauer, K., & Lewandowsky, S. (2018). Simple Measurement Models for Complex Working-Memory Tasks. Preprint. Retrieved from https://osf.io/vkhmu/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.

Sample Size Determination for the Bayesian t-test

Qianrao Fu, Herbert Hoijtink & Mirjam Moerbeek

Utrecht University

When two independent means are compared, H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 \neq \mu_2$, and H_2: $\mu_1 > \mu_2$ are the hypotheses of interest. This presentation introduces the R package SSDbain, which can be used to determine the sample size needed to evaluate these hypotheses using the Bayes factor. Both the Bayesian Student's t-test and the Bayesian Welch's t-test are available in this software package. The sample size is determined such that the median Bayes factor exceeds a user defined cut-off value. Topics that will receive attention are: SSD for H_0 versus an a priori point and an a priori distribution alternative; prior sensitivity; and, the use of Bayes factor as a measure of support and as a decision criterion. Using the R package SSDbain and/or the tables and figures provided in this presentation, psychological researchers can easily determine the required sample size. Statistical power in the null-hypothesis significance testing framework (NHST) has been studied for more than 50 years. Cohen played a pioneering role in the development of effect sizes and power analysis, and he provided mathematical equations for the relation between effect size, sample size, Type I error rate and power. However, the criticism with respect to the p-value is steadily increasing. BF for Bayesian hypothesis testing evaluation is increasingly receiving attention from psychological researches. Therefore, it is important to determine the sample size required before collecting the data for a Bayesian hypothesis testing. In classical power analysis for NHST, optimal sample size determination (SSD) is a means of choosing the smallest sample size to control the Type I and Type II error rates, and the relationship between sample size and power can be expressed by formulae. A simulation based approach will be used in this paper to

calculate the sample size needed for the Bayesian t-test to have sufficient support for the true hypothesis. To determine the sample size for a Bayesian evaluation of hypotheses with respect to two independent means the following ingredients are needed: (1) the choice for the Bayesian Student's t-test or Bayesian Welch's t-test; (2) a two-sided or a one-sided alternative hypothesis; (3) using a pre-specified effect size or a distribution of effect sizes under the alternative hypothesis; (4) decide what the desired support in terms of the median BF should be when either of H_0 and H_i $(i = 1, 2)$ is true. Then a series of results are shown. With the growing popularity of Bayesian statistics, it is important tools for sample size determination in the Bayesian framework become available. In this presentation, we develop software to calculate sample sizes within the framework of the Bayesian t-test hypothesis using time-efficient algorithms. In our future research, we will extend to more advanced statistical models, such as Bayesian ANOVA, ANCOVA, linear regression, and general multivariate SSD problems.

The analysis of the response profile of Motion and Form coherence tests by means of half normal psychophysical function

Sara Giovagnoli[a] , Roberto Bolzani[a] , Luca Mandolesi[a] , Kerstin Hellgren[b] , Sara Garofalo[a] & Mariagrazia Benassi[a]

[a] University of Bologna, [b] Karolinska Institutet

In literature, numerous studies investigated the role of vision in human cognitive development (Braddick & Atkinson, 2011). The functionality of the dorsal and ventral visual systems are usually evaluated with behavioral psychophysiological measures, such as motion perception (dorsal pathway) and form perception (ventral pathway). Such tasks use different coherence levels (signal-to-noise ratio) to evaluate subjective performance. However, a heated debate concerns the identification of the appropriate method for assessing the threshold. In particular, for the definition of the thresholds of motion and form coherence perception, the staircase adaptive procedure seems to be one of the most commonly used techniques (Ellemberg 2004; Armstrong & Maurer 2009; Hadad, Maurer & Lewis, 2011; Harvey, 1986). A different approach is represented by fitting the experimental data by a psychometric function (Parrish,

2005; Lewis, 2002). The use of an appropriate function allows defining the stimulus threshold as the level of the stimulus that leads to a preselected level of correct answer (e.g. 75%). All these methods are heterogeneous and not exempt of criticism. The aim of this study is to evaluate the applicability of a half-normal psychophysical function to fit the accuracy profiles obtained by Motion and Form coherence test. A sample of 48 adults (32F, mean age= 23.6, SD= 3.0) completed the Form and Motion coherence tests with 5 different coherence levels. The response profiles of the two tasks were fitted by the half-normal psychophysical function, to estimate the discrimination performance (i.e. the number of correct responses) related to the coherence level of the stimulus. The fitting function showed, in particular for the Motion coherence test, a quite good index of goodness of fit, suggesting the adequateness of half-normal cumulative function to represent motion coherence perception data. This method allows to characterize subjective performance with the W parameter, representing an estimation of the standard deviation of the distribution, and to compare the task difficulty of the Motion and Form coherence tests.

Incorrect responses in the response time interaction contrast

Matthias Gondan

University of Copenhagen

Townsend and Nozawa (1995, Journal of Mathematical Psychology) derived predictions for response time interaction contrasts that distinguish several classes of cognitive architectures (serial, parallel, coactive, exhaustive, self-terminating) in double factorial experiments. Their original theorems were limited to experimental tasks with ceiling accuracy. In this theoretical note I investigate systems factorial technology (SFT) within two canonical classes of models generating incorrect responses, namely, models with independent racers for informed responses and guesses, and models with mutually exclusive separate states for informed responding and guessing. I derive generalized interaction contrasts under these two model classes; these turn out to be related to the Kaplan-Meier and Aalen-Johansen estimators known from survival analysis. I discuss the limitations of the SFT approach if the incorrect responses arise from the component processes, and propose an alternative

experimental setup that varies the temporal onset of the stimulus components. I demonstrate that with onset delay, SFT methodology can be generalized to non-perfect accuracy, and I point out the consequences for response time experimentation.

Do Items Order? The Psychology of IRT Models

Julia M. Haaf

University of Amsterdam

Invariant item ordering refers to the statement that if an item is harder for one person it is also harder for everyone else. Whether invariant item ordering holds or not is a psychological statement because it describes how people may qualitatively vary. Yet, modern item response theory (IRT) makes an a priori commitment to item ordering. For example, the Rasch model limits items to invariant item ordering, and, conversely, the 2PL model prohibits items to order. Needed is an IRT model where invariant item ordering or its violation is a function of the data rather than an a priori commitment. For this purpose, a two-parameter shifted-exponential model is proposed that allows for the assessment of item ordering. Computational issues with shift-scale IRT models are discussed.

Quantum rotation: a new method for capturing a change of perspective

Thomas Hancock & Stephane Hess

University of Leeds

Quantum probability, first developed in theoretical physics, has also been used to model previously unexplainable cognitive data. This has led to the recent development of choice models based on quantum probability. Furthermore, quantum models can be used to accurately capture the ‘change of decision context and mental state’ when moving between choices made under revealed preference and stated preference settings. This paper tests whether these models can also capture ‘changing states’ or equivalently ‘changing perspectives’ in moral contexts. Under quantum models, each choice scenario is represented by a multidimensional ‘Hilbert’ space. If two choices are equivalent, then they can be represented by the same Hilbert space.

However, ‘incompatible’ choices are represented by different Hilbert spaces. To capture the change of perspective, a ‘quantum rotation’ is required, which maps the state vector from the basis of vectors representing the first choice, to the basis of vectors representing the second choice. Hancock et al., (2019) have previously demonstrated that this rotation accurately captures the difference between best and worst choice. In this paper, we use the same concept of quantum rotation to capture changes of context in moral choice scenarios. In our first dataset, decision-makers choose between the introduction of a new transport policy or keeping the status quo. A choice is defined as involving a ‘taboo trade-off’ if a decision-maker could choose a policy that involved decreasing tax or travel time at the cost of increasing the number of injuries or deaths. Here, a quantum rotation is used to capture the shift in perception of the alternatives in the presence of taboo trade-off. Allowing for a quantum rotation in such cases improves the log-likelihood of our quantum model from -725.4 to -716.9 (compared to an improvement from -721.2 to -719.5 in Chorus et al., (2018)’s model adding a penalty for taboo trade-offs).The second dataset we test involves decision-makers completing two sets of route choice tasks. The first set involved trade-offs between increased travel time and salaries for an individual, whereas the second additionally included attributes for increased travel time and salaries for the partner of the decision-maker. Crucially, including a quantum rotation for the change of basis when the decision-maker additionally considers attributes affecting their partner results in an improvement in log-likelihood from -12,656 to -12,348. Thus it appears that ‘quantum rotations’ accurately capture a change in decision context when a moral element enters the dimension of choice. Alongside these applications to moral decision making, the work also provides extensions to quantum models developed by Hancock et al., 2019, demonstrating that sociodemographics, mixed parameters, nesting structures, latent classes and error components can all be included within a quantum framework.

Response Time Extended Multinomial Processing Trees in R

Raphael Hartmann, Karl Christoph Klauer & Lea Johannsen

University of Freiburg

Estimating cognitive process completion times has been an active research area since Donders' "method of subtraction" (Donders, 1868). With the model class by Klauer and Kellen (2018) called Response Time extended Multinomial Processing Tree (RT-MPT) doing so is possible for every assumed process of a given MPT model. This model class overcomes some disadvantages of traditional MPT models. It can be used in many situations even when an MPT model is not identified, incorporating response times makes the probability estimates for MPT models more accurate, and variants of an MPT model can be tested against each other. So far, modeling RT-MPTs was only possible with C++. Therefore, we developed the R package "rtmpt" with which it is possible to fit RT-MPT models easily. The package is free and open source, it can be used with two established MPT syntaxes, and has a number of useful and new features such as suppressing specific process times, holding process probabilities constant, and changing some prior parameters. The package leads to parameter estimates that are comparable to the original C++ program by Klauer and Kellen (2018). Furthermore, we show that the Bayesian algorithm of the program is valid.

Bayesian Inference for Multinomial Models with Linear Inequality Constraints

Daniel W. Heck[a] & Clintin P. Davis-Stober[b]

[a] University of Mannheim, [b] University of Missouri

Many psychological theories can be operationalized as linear inequality constraints on the parameters of multinomial distributions. These constraints can be described in two equivalent ways: Either as the solution set to a system of linear inequalities or as the convex hull of a set of extremal points (vertices). For both representations, we describe a general Gibbs sampler for drawing posterior samples in order to carry out Bayesian analyses. We also summarize alternative sampling methods for estimating Bayes factors for these model representations using the encompassing Bayes factor method. We

introduce the R package multinomineq, which provides an easily-accessible interface to a computationally efficient implementation of these techniques.

Representing misconceptions within polytomous Knowledge Structure Theory

Jürgen Heller

University of Tübingen

Knowledge Structure Theory (KST) was developed for representing and assessing knowledge. In its original form it applies to a domain of dichotomous items (solved vs. not solved) and, within a competence-based extension, a set of dichotomous skills (present vs. absent) underlying the observed responses.The scope of KST has recently been widened by generalizing it to handle polytomous items with partially ordered response values (Heller, 2017). The talk presents a competence-based extension of polytomous KST, which is shown to provide a framework for an integrated representation of knowledge and misconceptions. The flexibility of this theory does not only allow for reconstructing previous approaches (de Chiusole et al., 2018; Lukas 1997), but offers the tools for formulating different assumptions about the structural relation between knowledge and misconceptions.

References:

de Chiusole, D., Gondan, M., Stefanutti, L. (2018). Probabilistic models for misconceptions in knowledge space theory. Paper presented at the *Meeting of the European Mathematical Psychology Group EMPG 2018*, University of Genova, Italy, 31 July 2018.

Heller, J. (2017). Knowledge space theory for polytomous items. Paper presented at the *MathPsych 2017*, University of Warwick, United Kingdom, 25 July 2017.

Lukas, J. (1997). Modellierung von Fehlkonzepten in einer algebraischen Wissensstruktur [Modeling misconceptions in analgebraic knowledge structure]. *Kognitionswissenschaft, 6*(4), 196-204.

Evidence accumulation in same-different judgments

Andrew Hendrickson[a] , Danielle Navarro[b] & Chris Donkin[b]

[a] Tilburg University, [b] University of New South Wales

Stimulus similarity plays a fundamental role in human cognition, shaping theoretical accounts of category learning, inductive reasoning, memory, and others. In this paper we introduce an evidence accumulation model for similarity based decisions that successfully accounts for the complete joint distribution over choice and response time across different stimuli, and how these distributions change systematically as a function of instructional demands. The modeling framework captures the way in which information about simple feature matches drives the early stages of stimulus processing, whereas the later stages are more heavily influenced by structural knowledge about the stimulus. Despite recent work showing that single process models are very often able to accommodate phenomena that ostensibly provide evidence for multiple processes, we show that no single process model provides a qualitatively reasonable fit to the results from two experiments.

Bayesian Evaluation of Informative Hypotheses

Herbert Hoijtink

Utrecht University

Since Cohen's (1994) paper "the earth is round, p< .05" there is increasing awareness that the null-hypothesis, e.g., H0: m1=m2=m3, where the m's denote the means in three groups, only rarely represents the expectations that researchers have. Informative hypotheses (Hoijtink et al., 2018) use equality and inequality constraints to formally represent researcher's expectation. Two (hypothetical) examples of such hypotheses are: H1: m1 > m2 > m3 and H2: m1 – m2 > m2 – m3. Since both H1 and H2 may be wrong, it is customary to add Hu: m1, m2, m3 to the set of hypotheses of interest. In Hu there are no restrictions on the parameters of interest. Only if H1 and H2 are better than Hu they may be valuable. Additionally, in the last years there is increasing awareness of the limitations and misuse of hypothesis evaluation by means of the p-value. An alternative, (informative) hypothesis evaluation using the

Bayes factor will be introduced. The Bayes factor (Gu et al., 2018; Hoijtink et al., 2018) quantifies the support in the data for a pair of hypotheses based on the fit and the complexity of the hypotheses. Loosely formulated, if, for example estimates of the three means in H1 are, 2, 5, and 7, respectively, then the fit of H1 is rather bad. It can also be seen that H1 is more specific than H2 (and therefore less complex) because it imposes more constraints on the three means. If, for example, BF12 = 5 and BF1u =10, this means that the support in the data for H1 is 5 times larger than the support for H2 and 10 times larger than for Hu. This would imply that, currently, H1 is the best available description of the population of interest. In the presentation it will be elaborated what the Bayes factor is, how it can be applied and should be interpreted. There will be attention for Bayesian updating (an alternative for power analysis), Bayesian (conditional) error probabilities, limitations of the approach, and the statistical underpinnings of the software with which the Bayes factor can be computed https://informative-hypotheses.sites.uu.nl/software/bain/

References:

Cohen, J. (1994). The earth is round, p< .05. *American Psychologist, 49*, 997-1003.

Gu, X., Mulder, J., and Hoijtink, H. (2018). Approximate adjusted fractional Bayes factors: A general method for testing informative hypotheses. *British Journal of Mathematical and Statistical Psychology, 71*, 229-261. DOI: 10.1111/bmsp.12110

Hoijtink, H., Gu, X., and Mulder, J. (2018). Bayesian Evaluation of Informative Hypotheses for Multiple Populations. *British Journal of Mathematical and Statistical Psychology*. DOI: 10.1111/bmsp.12145

Hoijtink, H., Mulder, J., van Lissa, C., and Gu, X. (2018). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*. DOI: 10.1037/met0000201

Towards meaningful inferences from attitudinal thermometer ratings

Yung-Fong Hsu[a], Michel Regenwetter[b] & James Kuklinski[b]

[a] National Taiwan University, [b] University of Illinois at Urbana-Champaign

Thermometer ratings and Likert scales are ubiquitous in social psychology, political psychology, and political science, even though critics have cautioned that researchers take the scores too literally. A measurement procedure based on arbitrary assumptions risks the real danger of generating scientifically meaningless inferences. Adopting a decision theoretic point of view, we use the concept of semiorders to capture the idea that a person giving two candidates distinct scores might or might not actually prefer one to the other, depending on the size of her threshold of discrimination. Furthermore, one respondent giving a candidate a lower score than another respondent could nevertheless be the stronger supporter. We state formal assumptions about the nature of preferences and propose a novel probabilistic response mechanism by which respondents construct numerical scores heterogeneously when asked to represent their preferences in a numerical format. We provide a proof of concept using maximum likelihood tests of our models on public domain American National Election Study data.

Diffusion models with time-dependent drift rates: A partial differential equation solution

Markus Janczyk[a] , Rolf Ulrich[b] & Thomas Richter[c]

[a] University of Bremen, [b] University of Tübingen, [c] University of Magdeburg

Diffusion models are nowadays used by psychological researchers to disentangle cognitive processes involved in decision tasks. The basic idea is that evidence for one or another decision is continuously accumulated until evidence exceeds a threshold and the corresponding response is emitted. At each time-step, a random amount of evidence is added, where the average amount depends on the drift rate. This accumulation process describes a Brownian motion with a drift. While the drift rate may vary between trials, it is

usually assumed to be constant within a trial, that is, not time-dependent. Besides Monte-Carlo simulations, fast solutions via a partial differential equation (PDE) approach are available for this type of model (Voss & Voss, 2008). In recent years, however, diffusion models have been advanced where the drift rate can vary within a trial, that is, the drift rate is assumed to be time-dependent. For example, the Diffusion Model for Conflict tasks (DMC; Ulrich et al., 2015) assumes that a temporary activation elicited by irrelevant stimulus features in conflict tasks, which rises fast and then declines. This brief activation is added to the activation of the controlled process that accumulates relevant information with a constant drift. The superposition of these two processes results in time-dependent changes of drift rate within a trial. To the best of our knowledge, no efficient PDE solution to this problem is available, but rather time-consuming Monte-Carlo simulations were required when fitting the model to data. Here we present a fast PDE solution implemented in Matlab and C++ that allows researchers to more efficiently fit diffusion models with a time-dependent drift rate to data. In this talk we introduce the problem and present several numerical solutions in comparison with the simulation approach.

Causal Interpretation of Statistical Models - Why we shouldn't ignore the scientific philosophers

Andreas Klein

Goethe University Frankfurt

The causal interpretation of statistical models and results has been a topic of interest both to statistical researchers and to the informed public who often expects statements about causal inferences in view of interventions in public health or economical domains. Statistical concepts and procedures, such as Rubin's causal model, have been recommended that claim to define causation and separate causal from pure correlational effects. Parallel to this, and not always received by those who intensively use math. models, there has been an ongoing debate in the philosophy of science literature about the challenges of a sufficiently precise and proper definition of causality (e.g., D. Lewis, P. Suppes, W. Salmon, H. Reichenbach), applying probabilistic or counterfactual concepts, or even drawing from modal logic. Most of them argue from an empiricist's perspective, and few

attempts have been made to delineate the implications for the real-world interpretation of math. models in psychology or the behavioral sciences. Are there consequences for covariate and variable selection, for the empirical significance of statistical quantities, or are there alternatives to modeling mean structures? This paper attempts to address some of these issues.

Alternatives to the Inverse Wishart distribution in Bayesian hierarchical IRT models

Christoph Koenig[a] *& Alexander Naumann*[b]

[a] Goethe University Frankfurt, [b] DIPF - Leibniz Institute for Research and Information in Education

The Inverse Wishart (IW) is one of the most popular prior distribution for covariance matrices in Bayesian hierarchical Item Response Theory (IRT) models. Its use, however, is problematic. First, it is not flexible because uncertainty for all variances is controlled by a single degree of freedom hyper-parameter ν. Second, when $\nu > 1$, the marginal distribution for the variances has a low density near zero. Third, there are a-priori dependencies between the correlation and the variance components. These characteristics make biased estimates of variance components and correlations, as well as item parameters and trait scores more likely. This simulation study compared the performance of three alternative specifications (scaled IW, hierarchical IW, and a separation strategy based on the Cauchy distribution and the Cholesky factor of the correlation matrix) to the standard IW in terms of parameter recovery (Bias and RMSE) in the hierarchical two-parameter logistic (2PL) IRT model. The design of the simulation consisted of four factors: sample size (N = 50, 100, 200), test length (k = 25, 50), the variance components of the item parameters ($t_{a,ß}$ = [0.2,0.5], [0.2,0.9], [0.6,0.9]), and the correlation between item parameters ($r_{aß}$ = 0.1, 0.3, 0.5). Results show a superior performance of the alternative specifications, especially of the separation strategy, regarding recovery of the variance components of the item discrimination parameter for smaller variances across all test lengths and correlations. For the larger variance component of the item discrimination, and in case of the

variance component of the item difficulty parameter, results indicate a better performance of the standard IW. This pattern can be explained, however, by the distinct peak and light tails of the marginal prior distribution the variance components, which yields more conservative estimates than the heavy-tailed Cauchy distribution. There were small advantages of the separation strategy over the other alternative distributions and the standard IW regarding recovery of the item discriminations. Differences in item difficulties and trait scores were negligible. The results of this simulation study will be discussed in the context of an optimized 2PL model for small sample item calibration and the question if the IW should retire from its service in Bayesian hierarchical IRT modeling.

Variance Constraints For Hierarchical Signal Detection Models

Martin Lages

University of Glasgow, University of Tübingen

Signal detection theory has a long tradition (Tanner & Swets, 1954; Green & Swets, 1966) and various models have been put forward to evaluate performance in detection and discrimination tasks (Macmillan & Creelman, 2005; Wickens, 2001). More recently, hierarchical SDT models have been suggested that can estimate sensitivity and response bias on individual and group levels (Rouder & Lu, 2005; Lee, 2008). These Bayesian models with equal or unequal Gaussian variances typically rely on ROC data and therefore require additional measurements from pay-off conditions, confidence ratings, or response times (Morey, Pratte & Rouder, 2008; Selker et al., 2019; Starns & Ratcliff, 2014). Here, we discuss an extension of hierarchical SDT models that, together with an informative prior, exploits variance constraints between signal and noise distributions. Variance-constrained SDT models were applied to data sets on inductive and deductive reasoning (Reit & Rotello, 2005). The results illustrate that these models can improve estimation of critical parameters. A constrained SDT model with unequal variance had lower deviance information criteria (DIC) and estimates comparable a parameter-expansion SDT model (Lee & Wagenmakers, 2013). A constrained SDT model with equal variance had the lowest DIC but gave slightly different estimates. Introducing variance constraints may enable researchers to improve data analysis of signal detection

experiments and to directly compare SDT models with equal and unequal variance.

Traps and tricks in Monte-Carlo simulation-based parameter estimation of advanced mathematical models

Yiqi Li, Martin Schlather & Edgar Erdfelder

University of Mannheim

Mathematical models applied in psychology and cognitive science have become more and more sophisticated, taking complex psychological processes and the interactions between them into account. For newly developed mathematical models, established methods for statistical inference that can be applied directly are not necessarily available. Often, the model cannot be described analytically so that numerical methods based on Monte-Carlo simulation are required to obtain model predictions such as the density function. Fitting such models can be challenging and the modeler may be confronted with a series of problems. These problems usually fall into two categories: extensive computation time and identifiability issues. Pursuing the goal of reducing computational intensiveness as well as numerically undesirable phenomena, we promote programming strategies and techniques to diagnose and resolve diverse kinds of problems. Regarding computation time, we demonstrate that some seeming trifles, such as the choice of the programming language, the design of the simulation study and the storage and management of intermediate results, can actually have a significant influence on the speed. In contrast, parallel computing and vectorization do not always yield a gain, depending on the modeling problem and the implementation. Low computational expense facilitates the detection and rectification of identifiability issues, such as ill-conditioning, correlations among parameters, and over-parameterization. We show that numerical undesirable characteristics can arise artifactually from unsuitable use of random seeds and the way in which random numbers are generated and assigned. We also discuss how reparameterization and profiling can yield proper box constraints, solve convergence problems, and reduce dependency of the results on starting values. Furthermore, we present an algorithmic procedure for the case of fitting a model to multiple datasets collected under different

experimental conditions, whereby some of the parameters are assumed to depend on these conditions while others to be invariant to them. We explain our approaches using the example of a queueing model of visual search (Li, 2018). In conclusion, an optimal application of carefully chosen numerical and programming strategies and techniques can speed up the parameter estimation process by a factor of 100 and lead to more stable and more accurate estimates.

Reference:

Li, Y. (2018). Markovian queueing model of visual search with integrated error analysis. Talk at the *49th Meeting of the European Mathematical Psychology Group*, Genova, Italy.

Testing the k-modal race model inequality

Luigi Lombardi[a] *& Hans Colonius*[b]

[a] University of Trento, [b] University of Oldenburg

The race model inequality (RMI) implies an upper bound on the amount of statistical facilitation for reaction times (RTs) attainable by a race model in redundant-signals tasks. Over the years, it has become a relevant and popular testing tool to measure the amount of statistical evidence against a race model. Here we extend a recent nonparametric procedure (for single participant analysis) to evaluate the RMI in the two-modal representation (Lombardi, D'Alessandro, & Colonius, 2018) to the more general k-modal case (with $k \geq 2$) in a group analysis context. In particular, the generalization of the truncated property of the reconstructed distribution under maximal statistical facilitation for a race model is illustrated together with some technical connections with the Vincentizing (quantile averaging) procedure. Some simulation results suggest that our statistical test efficiently controls for type I error with adequate power.

Introduction of right action on extensive structures for intertemporal choice

Yutaka Matsushita

Kanazawa Institute of Technology

The aim of this study is to introduce a right action of positive real numbers on extensive structures. So far, focusing on a generalized extensive structure as a base space, we have considered introducing (Matsushita, 2017) an action of right multiplication of subjective time duration. The generalized extensive structure has a special element e, a left identity, that plays an important role in expressing the advanced or postponed receipt of outcomes. The right multiplication (resp. right division) of an outcome a by e indicates advancing (resp. postponing) its receipt by one period. Meanwhile, the right multiplication or right division by any positive real number is also provided on the generalized extensive structure to express the advancement or postponement by arbitrary subjective time duration. This causes a problem with the interpretation of the time duration (i.e., real number) corresponds to e. It is natural to identify e with the null time duration 0. It is also known (Matsushita, 2014) that the generalized extensive structure reduces to the original extensive structure whenever e is a two-sided identity. Consequently, by identifying e with 0, all properties that are satisfied for the generalized extensive structure are satisfied for the original extensive structure. This identification not only solves the above-mentioned problem but also brings us the merits. First, it makes homogeneity of right action a trivial axiom. Second, it revokes an artificial axiom (commutativity of multiplication between e and any positive real number). The effect of the postponement or advancement is discussed in detail further. It should be noted here that a/t is defined as an outcome received at the present that is equivalent to the outcome a at the postponed time t. Let t_i be an increment in a subjective time duration when the i-th postponement occurs, and let T be the total time duration calculated by concatenating all increments t_i. In general, $(\cdots((a/t_1)/t_2)\cdots)/t_n$ is not equivalent to a/T, where the left side implies the n-step postponement accompanied by t_i, and the right side implies the one-step postponement by the total duration T. The n-step postponement may or may not reduce the value of a more than the one-step postponement, even though the total postponement durations are the

same. This is attributed to our basic concept that the decrease of the value of an outcome, received at some delayed time, by a further delay can also depend on the time. In addition, it is shown that if the step-by-step postponement has no effect (i.e., if the left and right sides are equivalent), then the reciprocal of the weight function of our utility model becomes an exponential discount function.

Learning to Compare

Louis Narens

University of California

In interactive situations, agents can ‘‘learn’’ something that is not a preexisting truth. They can converge to an arbitrary convention, or tacit agreement. Once established they may even view it as an objective truth. Here we investigate accommodation dynamics for interpersonal comparisons of utility intervals. We show, for a large class of dynamics, convergence to a convention.

Bayesian analysis of information used during decision making

Tillmann Nett[a], Nadine Nett[a] & Andreas Glöckner[b]

[a] FernUniversität in Hagen, [b] University of Cologne

In research on decision making, some models assume that parts of the provided information are ignored. For example, the take-the-best (TTB) heuristic assumes that only the most valid cue is used. Other findings indicate that even irrelevant information is used. For example, when the options are stereotypically male or female objects, the gender of the experts providing the recommendations is used as well. Unfortunately, currently, testing which information is used also requires forming hypotheses about how this information is integrated with other pieces of information. We show, that in experiments with two options different information usage leads to different equivalence classes over the set of trials. Intuitively, two trials must appear the same, if they only differ in regard to information which is ignored. Similarly, two trials can appear to be mirror images of each other under reduced information, when the retained information is mirrored. This allows us to define a formal structure that we call perceptive frame, which describes the cognitive filter that is applied

during decision making. Furthermore, for an experiment with N trials, the perceptive frames have a simple geometric structure corresponding to subspaces of the N-dimensional vector space. The elements of these subspaces can be shown to corresponds to hypotheses about experimental outcomes expressed in log-odds for choosing one of the options. Thus, using a basis transformation perceptive frames allow for a simple parametrization that can be used to define closed formulas for the Bayes-factors. Additionally, the formal structure allows us to identify a partial ordering over the set of perceptive frames. In this ordering, a perceptive frame is related to all other perceptive frames, which use at least the same information. The resulting partially ordered set (poset) is categorically equivalent to a thin category of enumerating functions from the set of trials to the whole numbers. By showing that the product of any two objects exists in this category, we also prove that the perceptive frame poset has the algebraic structure of a join semi-lattice. Since the product has a simple implementation, this allows us to easily combine any two perceptive frames into another perceptive frame that encompasses the information used in both original perceptive frames. Thus, our method only requires the definition of a few very simple perceptive frames, while the rest of the semi-lattice can be completed automatically. Furthermore, subsets of the perceptive frames corresponding to an experimental hypothesis can be selected automatically from the completed semi-lattice based on the partial ordering. We test our method in a simulation study and show that the retained information can be identified with high accuracy (area under ROC curve > .92). Additionally, we provide some examples of how this method can be applied to experiments.

Dynamic Choice with Status Quo: Theory and Design of Efficient Experiment

Hassan Nosratabadi & Francois Maniquet

Université catholique de Louvain

We study a model of decision making in which choice situations come in a specific order and the decision maker displays the following status quo bias: she sticks to her previous choice whenever it is still available, except if a new alternative is obviously better. As a result, only changes in the chosen alternative along the choice

sequence reveal preference. We show that the resulting dynamic status quo biased choice model is characterized by the strong axiom of revealed preferences along with the weak axiom of revealed obvious preferences. Finally, we show that, provided the number of alternatives is sufficiently large (at least 5 alternatives), the length of the minimal sequence of choice situations needed for extracting unique preferences is not larger under status quo bias than under rationality.

A graphical taxonomy of assessment models

Stefano Noventa & Jürgen Heller

University of Tübingen

In the past years, several theories and frameworks for assessment have been developed within the overlapping fields of Psychometrics and Mathematical Psychology. The most notable are Item Response Theory (IRT) models, Cognitive Diagnostic Models (CDMs), and Knowledge Space Theory (KST). Yet, in spite of their common goals, these theories have been developed largely independently, and focus on slightly different aspects. In spite of the methodological differences, these theories exhibit several similarities, and various attempts to bridge them can be found in literature. To name only a few, connections between CDM and KST (Heller et al., 2015), between KST and IRT (Stefanutti, 2006; Noventa et al., 2018), or between CDM and IRT (Junker & Sijtsma, 2001; Hong et al., 2015) have been highlighted. A particularly interesting similarity lies in the treatment of their conditional probability parameters (i.e., lucky guesses and careless errors in KST, slipping and guessing parameters in CDM, and guessing and ceiling parameters in IRT). In order to capture this structural similarities, a two-processes model is introduced. It separates guessing and ceiling parameters into a first process modeling the effects of individual ability on item mastering, and a second process representing the effects of pure chance on item solving. Based on this model, a graphical taxonomy encompassing IRT models, CDMs, and KST models is thus obtained. Some consequences for both dichotomous and polytomous items are discussed.

The Strategy Aggregation Effect in Group Judgment

Henrik Olsson & Mirta Galesic

Santa Fe Institute

What characteristics of cognitive strategies affect the predictive accuracy of individual and group judgments? We relate the literature on model performance in statistics and machine learning to performance of realistic cognitive strategies in individual and group settings. We investigate two well-studied classes of cognitive strategies: unconstrained and constrained linear judgment strategies. We show that constrained strategies are more accurate for individual judgments, but when individual judgments are aggregated to produce a group judgment an unconstrained strategy is more accurate. This strategy aggregation effect can be understood by analyzing a decomposition of the mean squared error into bias, variance, and covariance. Because of their lower bias but higher variance, unconstrained strategies perform worse for individual judgments, but better for group judgments where aggregation minimizes variance. In computer simulations with artificially constructed and real environments, we further show that this aggregation effect does not occur if there are correlations between individual judgments. Here, constrained strategies always outperform an unconstrained strategy, because the larger covariance component of the unconstrained strategy outweighs its lower bias.

Diffusion models with time-dependent drift rates: Numerical accuracy and efficiency in simulation and parameter estimation

Thomas Richter[a] , Markus Janczyk[b] & Rolf Ulrich[c]

[a] University of Magdeburg, [b] University of Bremen, [c] University of Tübingen

Diffusion models are based on the assumption that the decision process is continuous in time. Numerical evaluations of these models, however, involve discrete time steps. This discretization of time implies a non-exact approximation of the diffusion process. Numerical evaluations are based on one of the following three approaches: (a) Monte-Carlo simulations involving stochastic differential equations that, besides the time step size, require the

number of repeated random samples as an additional parameter; (b) analytical solutions that are based on the artificial truncation of infinite sums representing the distribution function; (c) a partial differential equation (PDE) solution, namely the Kolmogorov backward equation, whose evaluation requires a further spatial discretization parameter. The exact diffusion process is obtained if these parameters reach a limit, that is, time step and spatial discretization parameter approach zero, and the number of repeated random samples in stochastic simulations or terms included in the analytical sum converge to infinity. However, this ideal setting cannot be realized and hence numerical approximation errors are inevitable. Consequently, a feasible tradeoff between computational accuracy and computation time is required. In this talk, we examine the Monte-Carlo approach and our PDE solution for simulating decision processes with a focus on non-standard settings including time-dependent drift rates and time-dependent thresholds. Furthermore, we analyze the role of various approximation parameters like time step size, spatial step size, and the number of repeated random samples. Finally, we discuss the relevance of numerical accuracy when it comes to fitting diffusion models with a time-dependent drift to observed data.

Understanding Belief Polarization - An Agent-Based Modeling Approach

Nadia Said[a], Debora Fieberg[a], Helen Fischer[a], Andreas Potschka[a] & Christian Kirches[b]

[a] Heidelberg University, [b] TU Braunschweig

We implemented an agent-based model (ABM) to explore the influence of cognitive parameters, like working memory capacity as well as openness, on polarization of beliefs in an initially heterogeneous belief environment and showed how individual differences in openness and working memory capacity can accelerate belief polarization. In this contribution we will present an extended version of our previous model and discuss (a) the influence of different initial distributions on belief polarization, (b) explore how bots influence the propagation of beliefs, (c) show simulation runs with real data from our experiment. Furthermore, we will outline the setup of our mathematical model of the ACT-R declarative memory

module (Said et al. 2016) that will allow us to simulate the cognitive process of each agent in a more realistic way.

Efficient Hypothesis Tests in Multinomial Processing Tree Models: A Sequential Probability Ratio Test for the Randomized Response Technique

Martin Schnuerch & Edgar Erdfelder

University of Mannheim

In a seminal article, Riefer and Batchelder (1988) proposed Multinomial Processing Tree (MPT) models to measure latent psychological attributes based on categorical behavioral data. Ever since, numerous MPT models have been developed and successfully applied in different areas of psychological research. One class of these models aims at overcoming response biases to sensitive questions such as “Have you ever taken cocaine?”: The Randomized Response Technique (RRT) protects individual respondents’ privacy by prompting them to adjust their answers according to the outcome of a randomization device. To allow for estimation of the unknown prevalence of the sensitive attribute, RRT models incorporate parameters for the prevalence and the known randomization probability. The technique has repeatedly been shown to produce more valid estimates than direct-questioning formats. Due to randomization, however, RRT parameter tests typically have low statistical power, often resulting in large required sample sizes when testing hypotheses on the prevalence of a sensitive attribute. As a remedy, we propose a Sequential Probability Ratio Test (SPRT) for RRT models. In contrast to traditional analyses for fixed sample sizes, sequential statistical procedures continuously monitor the data during the sampling process and terminate when a predefined decision criterion is met. Sequential analysis may thus substantially reduce the required sample size without increasing long-term error rates. We show how to implement the SPRT for common RRT variants in standard statistical software. Moreover, we demonstrate analytically and by means of simulations that the SPRT requires approximately 50% smaller samples for RRT models than traditional analyses. Finally, we illustrate the efficiency of the proposed sequential RRT using an empirical example.

Thinking fast, not slow: A drift diffusion model account of belief bias

Anna-Lena Schubert[a], Mário B. Ferreira[b] , André Mata[b] & Ben Riemenschneider[a]

[a] Heidelberg University, [b] Universidade de Lisboa

A widespread assumption of dual process models is that sound reasoning relies on slow, careful reflection, whereas biased reasoning is based on fast intuition. However, recent experimental results challenge a clear distinction between fast belief-based and slow rule-based processing. Instead, they suggest that rule- and belief-related problem features are processed in parallel and that problem features such as feature complexity determine whether participants give a rule- or belief-based response. In particular, a divergence between rule- and belief-based processing is supposed to cause mutual interference, resulting in lower accuracies and higher reaction times (Handley & Trippas, 2015; Trippas, Thompson, & Handley, 2017). We used the diffusion model to mathematically describe the competition between rule- and belief-based processes and directly test key predictions of the parallel process model in a study with N = 40 participants, who completed a syllogistic reasoning task in addition to cognitive abilities measures and personality questionnaires. Consistent with the parallel process account of belief bias, drift rates were smaller when rule- and belief-based problem features conflicted than when they aligned, whereas we found insufficient evidence for a criterion shift between conflict and non-conflict problems. Moreover, we found dissociations in the way drift rates related to individual differences that may be accounted for in terms of Stanovich's (2009) concepts of algorithmic and reflective thinking. While individuals with higher reflective thinking (as assessed by the cognitive reflection test and need for cognition) showed higher drift rates specifically in conflict trials, individuals with higher algorithmic ability (as assessed by working memory capacity) showed a greater general efficiency of information processing as reflected in higher drift rates in both conflict and non-conflict trials. In conclusion, our results suggest that more reflective reasoners inhibit interfering belief-based information better and process information more efficiently than biased responders. In this sense, they challenge the widespread assumption that sound reasoning depends on slow and cautious analytical processing.

An updated concept of evidence based on Bayes' law, which explains decision making for sensory tasks with numerous complicated objects

Valentin Shendyapin & Irina Skotnikova

Russian Academy of Sciences, Moscow

Bayesian inference (Knill, Richards, 1996; Kersten, Yuille, 2003) is used for modeling of sensory tasks performing in which an observer has to give a response: which of 2 alternative events (A or B) has been presented. According to the Bayes theorem, a posteriori probabilities of the events A and B are calculated as:

$$P(A|\boldsymbol{x}) = [P(A)f(\boldsymbol{x}|A)] / [P(A)f(\boldsymbol{x}|A) + P(B)f(\boldsymbol{x}|B)], \quad (1)$$

$$P(B|\boldsymbol{x}) = [P(B)f(\boldsymbol{x}|B)] / [P(A)f(\boldsymbol{x}|A) + P(B)f(\boldsymbol{x}|B)], \quad (2)$$

where $P(A)$, $P(B)$ are a priori probabilities of A and B; $f(\boldsymbol{x}|A)$, $f(\boldsymbol{x}|B)$ are a priori distributions of probability densities of sensory effects vector $\boldsymbol{x}$. The decision rule for a response is:

an observer met the event A if $P(A|\boldsymbol{x}) > P(B|\boldsymbol{x})$, otherwise he (she) met B. (3)

In our model describing discrimination between similar stimuli A and B (Shendyapin, Skotnikova, 2011) we have introduced a new variable:

$$\Psi = ln[P(A) / P(B)] + ln[f(\boldsymbol{x}|A) / f(\boldsymbol{x}|B)]. \quad (4)$$

It has allowed us to rewrite probabilities (1, 2) as:

$$P(A|\boldsymbol{x}) = P(A|\Psi) = 0.5[1 + th(\Psi/2)], \quad (5)$$

$$P(B|\boldsymbol{x}) = P(B|\Psi) = 0.5[1 - th(\Psi/2)]. \quad (6)$$

If the variable Ψ increases from $-\infty$ to $+\infty$, then the probability $P(A|\Psi)$ increases monotonically from 0 to 1. Therefore $P(A|\boldsymbol{x})$-value can be measured in units of Ψ. Since $P(A|\boldsymbol{x})$ becomes greater along with greater Ψ, then Ψ is "an evidence in favor of A". On the evidence axis Ψ, there is a single point $\Psi_{cr} = 0$, where the probabilities of 2 alternative events are equal to each other:

$P(A|\Psi) = P(B|\Psi) = 0.5$. The decision rule of our model does not require the both probabilities calculation. In order to give a response, it is enough to compare the evidence Ψ (obtained by the observer) and the criterion $\Psi_{cr} = 0$: if $\Psi > 0$ then A occurred, otherwise B occurred. (7)

Since the criterion $\Psi_{cr} = 0$ does not change during the observations, then the decision rule (7), which requires recalculation of only one Ψ-value for each response, is more convenient for realization than the rule (3), which requires recalculating of the two probabilities $P(A|x)$ and $P(B|\boldsymbol{x})$ each time. The computation of Ψ is much simpler than computation of $P(A|x)$ and $P(B|\boldsymbol{x})$, since it does not require multiplication and division of probabilities, but is reduced just to addition and subtraction of the logarithms of probabilities. And the operation of logarithm is not difficult for brain neurons.

Since the Bayes formula is applicable for not only two events but for any number of events, then the evidence Ψ may be used to scale a posteriori probabilities of any number of sensory events. Without loss of generality of inference, one can show it for three events. For each observation we may use formulas similar to (4) to introduce evidences Ψ_A, Ψ_B, Ψ_C in favor of each of the three alternative events A, B, C so that they satisfy the modified Bayes formulas: $P(A|\Psi_A) = 0.5[1 + th(\Psi_A/2)]$; $P(B|\Psi_B) = 0.5[1 + th(\Psi_B/2)]$; $P(C|\Psi_C) = 0.5[1 + th(\Psi_C/2)]$. In this case, the decision rule using to give a response is: if $\Psi_A, > \Psi_B, \Psi_C$, then the event A occurred; if $\Psi_B, > \Psi_A, \Psi_C$, then B occurred; if $\Psi_C, > \Psi_B, \Psi_A$, then C occurred.

Empirical Distinctness of Skill Map Based Knowledge Structures

Andrea Spoto & Luca Stefanutti

University of Padua

The identifiability of the BLIM have been recently investigated in several articles (e.g., Doignon, Heller, & Stefanutti, 2018; Heller, 2017; Spoto, Stefanutti & Vidotto, 2013; Stefanutti & Spoto 2018). These studies particularly focused on the relationship between parameter identifiability and certain properties of the knowledge structures, namely the presence of the so-called forward and backward gradedness. Furthermore, Spoto, Stefanutti & Vidotto

(2012) established some connections between the forward and backward gradedness of a knowledge structure and the properties of the skill map delineating it. In this research some necessary and sufficient conditions are specified which characterize a skill map delineating a forward and/or backward graded structure. Moreover, based on the two notions of forward and backward graded knowledge structures, the present research will reveal the existence of a more critical and severe type of "identifiability" of the BLIM. It is called "empirical distinctness" and it does not only involve the parameters of the probabilistic models, but also the deterministic and combinatorial structures on which they are based. We define two probabilistic knowledge structures $(Q, \mathcal{K}_1, \theta_1)$and $(Q, \mathcal{K}_2, \theta_2)$ as "empirically indistinguishable" if the prediction function of the BLIM maps both $\theta_1$and θ_2 to the same probability distribution on the response patterns. The empirical indistinctness concerns especially the two structures $\mathcal{K}_1$ and $\mathcal{K}_2$, besides the two parameter vectors. In fact, if $\mathcal{K}_1$ and $\mathcal{K}_2$ are empirically indistinguishable, then they are either both accepted, or both rejected, in every possible experiment or empirical data set. It will be highlighted how empirical distinctness goes above and beyond identifiability of the models. In fact, an identifiable structure could still be empirically indistinguishable from another one. It will be shown how empirical distinctness transfers to the skill maps delineating two (or more) indistinguishable structures. The consequences of both structures and skill maps indistinctness will be explored in terms of the possibility of uniquely either selecting the "true" structure or establishing the existence of skills underlying the responses to a set of items. Finally, it will be pointed out how the obtained results easily apply to similar frameworks, such as Cognitive Diagnostic Models.

Methodological flexibility in prior elicitation and its effects on Bayesian model comparison

Angelika Stefan

University of Amsterdam

Bayesian model comparison is a versatile approach to contrasting the psychological hypotheses contained within mathematical models. A fundamental part of Bayesian model comparison is the precise definition of the selected models, which includes the specification of

prior distributions on parameter values. While some authors argue for default solutions, a growing number of researchers have advocated for informative priors, which provide more tightly constrained models based on the plausibility of different parameter values. Several different approaches have been proposed to create these informed priors, with the shape of the distribution being based on psychological theory, general logical considerations, or previous data. One of these approaches is prior elicitation. Prior elicitation aims to transform qualitative plausibility judgments of one or more experts into a quantitative probability distribution. Although elicitation is widely regarded as a best practice approach for prior specification, even the most robust elicitation procedures contain many potential sources of methodological flexibility. In this talk, we will discuss how this methodological flexibility can influence the elicited prior distributions, and therefore, the results of a Bayesian model comparison. For example, we will demonstrate how initial disagreement among experts, the choice of a specific elicitation technique, and different strategies for the mathematical combination of several elicitation results can each determine the outcome of the elicitation process. As the incorporation of expert knowledge in prior distributions is becoming an increasingly popular approach to solving the ubiquitous need for specifying prior distributions in Bayesian model comparison, we argue that more robust approaches and guidelines for the process of prior elicitation should be developed.

Extracting quasi-orders from polytomous data by a minimum discrepancy approach

Luca Stefanutti

University of Padua

Initially, applications of the theory of knowledge structures were only possible with items having a dichotomous response format. Later, an extension to items with more than two response alternatives was proposed by Schrepp (1997). However, one of the major difficulties with such an extension was how polytomous structures or spaces could be specified or constructed in practice. Recently, a k-medians algorithm which extracts polytomous structures from data was proposed by de Chiusole, Spoto, and Stefanutti (in press). It poses no restrictions on the type of polytomous structures that can be

extracted from data. However, sometimes the objective is to extract structures having specific properties like, for instance, respecting a certain order on the set of items. A new procedure is proposed in this talk, which extracts quasi-orders from both dichotomous and polytomous data, by using a minimum discrepancy approach. The procedure has the features of a clustering algorithm in the “response pattern classification” stage, but it differs from standard clustering because it has no cluster centroid updating stage. Let Q be a finite and nonempty set of items and L be a finite and nonempty set of linearly ordered “levels”. A polytomous state is a mapping K from Q to L. Such a mapping induces a weak order W on Q in the sense that pWq if and only if K(q) <= K(p). Then we say that K is consistent with a given quasi-order R if R is a subset of W. The collection of all the polytomous states consistent with R is named the polytomous space derived from R. Since a bijective correspondence exists between quasi-orders on a set of items, and the family of all the polytomous spaces on Q and L (see, e.g., Schrepp 1997), the discrepancy of a weak order from a given polytomous data set is nothing else than the discrepancy of the corresponding polytomous space from the data set. If d is a metric on L^Q, then a minimum discrepancy from the data can be obtained for every state in the polytomous space, and the problem is to find the quasi-order on Q that minimizes the sum of such minimum discrepancies. An efficient method for accomplishing this task is presented. Results of a simulation study and of two empirical applications to existing polytomous data sets are presented.

How to deal with rational intransitive choices

Reinhard Suck

University of Osnabrueck

When dealing with preferential choice intransitive choice behavior is generally considered an error of the deciding subject. However, there are several cases where it is completely rational. We investigate how two approaches to model choice behavior --- random utility theory and choice function theory --- can handle such situations. In the course of the investigation it turns out that choice functions can be described in the framework of random utility theory. Since the latter is closely connected to polytopes, mostly the linear ordering

polytope, it comes as no surprise that choice functions can be modeled probabilistically by drawing on the weak order polytope. The investigation raises doubts about the usefulness of the concept of rationalizability of choice functions. This concept is widely utilized and has born mathematically beautiful results on choice functions. The paper explores the tenability of the concept and how it can be modified.

Axiomatic properties of bad decision

Kazuhisa Takemura

Waseda University

We make decisions on every occasion: from casual ones such as what to eat for lunch to more serious decisions such as an individual's future course and government policy. We also sometime make bad decision. Bad decision can be operationally defined as choosing the worst alternative. Let X be a finite set to be selected. If the relation R is complete and acyclic, then the choice function of X holding R relation, C(R, X), is not empty. That is, under this condition, the worst option as well as best option exists. Let R be complete. Then, R is acyclic, if and only if C (R, X), which is the bad choice function of X with finite elements, is not empty. That is, under R with completeness, that R is acyclic is the necessary and sufficient condition that the choice function leads the worst option as well as the best option. From re-interpretation of Arrow's (1951) theorem, I also exemplify that composing a multi-attribute value function that satisfies all the conditions below is impossible for making bad decision as well as making good decision, meaning that conditions to satisfy connectivity and transitivity, which are the conditions for rationality, and the following conditions presumably appropriate for rational decision-making do not hold simultaneously. If multi-attribute decision-making has transitivity and connectivity properties for selecting bad choice, then these properties contradict with combination of no limitation of space for multi attribute decision making problems, Pareto principle, independence of irrelevant alternatives, and multiple attribute. In other words, weak order preference, no limitation of space for multi attribute decision making problems (U), Pareto principle (P), and independence of irrelevant alternatives (I) leads to single attribute decision making for bad

decision as well as good decision. Any multi-attribute choice function generating a quasi-transitive R and satisfying condition U, P, and I must be oligarchic for determining the bad decision as well as the good decision. I will further discuss on the perspectives of bad decision from the point of view of behavioral decision theory (Takemura, 2014, Behavioral decision theory, Springer; Takemura in press. Escaping from bad decision. Elsevier).

Factorial invariance and orthogonal rotation

Silvia Testa & Luca Ricolfi

University of Torino

In factor analysis, a common viewpoint posits that the main purpose of rotation (orthogonal or oblique) is to facilitate the interpretation of factors by the transformation of the initial solution (i.e. a particular loading matrix) into a simpler solution. However, according to Thurstone (1935, 1947) and Kaiser (1958), the simplification produced by a rotation is not an end in itself but a means to ensuring the formal property called "factorial invariance". In essence, factorial invariance means the stability of the loadings when the same tests (or items) are included in different batteries measuring the same latent variables and performed on the same population. Stated in other words, what the principle of factorial invariance requires is a kind of insensitivity to the mix. Unfortunately, the rotation methods suggested by Thurstone (1947) and by other researchers after him do not guarantee factorial invariance in its strict definition, even in cases where the requirements of the simple structure are substantially fulfilled. To the best of our knowledge, the only exception is the Varimax method in the special condition described in Kaiser (1958), i.e. when there are only two factors and the tests are co-aligned in the two-dimensional factor space. In this study, the principle of factorial invariance is investigated in two conditions that are more general than the one presented in Kaiser (1958), but not so general as advocated by Thurstone (1947). In the first condition (Length test), two test batteries are made of the same set of tests and differ only for the number of times the whole set of tests is replicated. In the second condition (Mix test), the two batteries are made of the same set of pure indicators (each test is a marker of a single factor), and they only differ for the mix of tests. For both conditions, 108 pairs of

loading matrices (corresponding to the initial factor solutions) were simulated varying the number of tests, factors, length or mix and subjected to the principal orthogonal rotation methods available in literature (Browne, 2010). The aim of the study is to evaluate which methods among those implemented in the GPArotation package of R (11 methods) were capable of passing the two tests, i.e. which of them produce the same rotated solution for the common part of the two batteries for each of the 108 pairs of matrices tested. The root mean square deviation (RMSD) was used as a measure of discrepancy between the rotated common parts of the two batteries. Given a rotation method and 108 matrix pairs, the test (Length or Mix) was considered to be passed if max(RMSD) < 0.00001.As a result, the Length test was easier to be passed than the Mix test. Only 3 out of 6 methods among the Orthomax/Crawford-Ferguson family (Quartimax, Biquartimax and Varimax) and 2 out of 5 non-classical methods (Entropy and Infomax) were successful on both tests.

A Lévy-Flight Model of Decision Making

Andreas Voss & Marie Wieschen

Heidelberg University

In psychology, decision making is often – and successfully – modelled with the diffusion model, which is based on the assumption that evidence accumulation follows a Wiener diffusion process, that is, evidence accumulates over time with a constant drift and normal noise. Here, I will present a model suggesting that noise in evidence accumulation is not Gaussian but is better described by heavy-tailed distributions. Thus, the evidence accumulation process is mapped no longer by a diffusion process but by a so-called Lévy-flight. An important characteristic of Lévy-flights is the incorporation of jumps in the process. In decision making, such jumps indicate sudden changes in the subjective believes about the current situation. In the present talk I will (a) discuss possibilities to estimate parameters of the Lévy-Flight model, (b) compare the fit of the standard diffusion model and the Lévy-Flight model to empirical data, and (c) present first evidence of both individual-related and task-related predictors of the "heavy-tailed-ness" of the noise distribution.

A computational cognitive process model for best-worst choice

Lena Wollschlaeger & Adele Diederich

Jacobs University Bremen

The simple (2N-ary) choice tree model (Wollschlaeger & Diederich, 2012, 2017) is a computational cognitive process model for preferential choice between multiple alternatives. It assumes that the decision maker sequentially compares pairs of attribute values between alternatives and within attributes. The resulting evidence is accumulated in two counters per alternative, whereof one counts positive evidence and the other counts negative evidence. The difference of the states of the two counters is called preference state and compared to two thresholds, a positive choice threshold, and a negative elimination threshold. Whenever a preference state hits one of the thresholds, the respective alternative is chosen or eliminated from the choice set. The most preferred alternative is either the first to be chosen or the last to be eliminated from the choice set. With its two thresholds per alternative, the simple choice tree model naturally extends to best-worst choice, that is, situations where the decision maker is asked to choose the most preferred and the least preferred alternative from a set. We show how the simple choice tree model accounts for best-worst choice data (including choice response times) and fit it to data from three experiments reported by Hawkins, Marley, Heathcote, Flynn, Louviere, and Brown (2013, 2014).

Posters

A Sequential Sampling Model for Visual Search RT Distributions

Steven Blurton

University of Copenhagen

We propose and test a TVA-based RT model adapted to account for RT distributions obtained in classic visual search tasks. The model is based on a random walk representing the difference between two Poisson counters, each representing evidence in favor of a "target present" or "target absent" response. It is further based on the assumption of parallel processing with limited visual processing capacity. In the feature search task, filtering by colour is highly effective when a target is present. By contrast, in the target absent condition, filtering is not possible. Both conditions are well described by the Poisson random walk model. In the conjunction search task, filtering by colour is also effective, but only affects the processing speed of about half of the distractors. Filtering by orientation, on the other hand, is less effective, making the task considerably more difficult. As an alternative, we tested a variant of the model in which groups of items are searched sequentially. We demonstrate that the Poisson random walk model is a plausible account of RT data obtained in feature search and conjunction search. The model predicts the RT distributions of correct responses and the overall response accuracy. The medians of incorrect responses are also well explained overall, but the low number of incorrect responses precludes any strong conclusions from error RT distributions. These results suggest that a relatively simple model can explain RT distributions in visual search. Following from this, data from other tasks are needed to justify more elaborate and complex models.

A problem space approach to studying strategic planning in board games

Andrea Brancaccio & Luca Stefanutti

University of Padua

A link between Knowledge Space Theory (Doignon & Falmagne, 1995) and Problem Space Theory (Newell & Simon, 1972) was recently established, and it led to the creation of a procedure to obtain a knowledge structure from a problem space (Stefanutti, 2018). The present work extends this link, introducing the definition of homomorphisms between problem spaces. As a consequence, it is possible to establish an order relation between two problem spaces if it is possible to define a homomorphism between them. Such extension is used to set a framework to obtain a knowledge structure from a board game-derived problem space. Board games are complex domains of knowledge, and a problem space for a board game is usually huge and difficult to manage. Redefining operations and problem states in order to obtain and new problem space which is homomorphic to the original one, allows us to obtain smaller and more manageable problem spaces and consequently smaller and more manageable knowledge structures as well, while preserving important properties of the original problem space like the sub-problem relation. Board games also involve two or more players, and to manage the presence of two problem solvers with different goals, a minimax criterion was defined to obtain the solution paths for each problem. Depending on the problem solver playing just as himself or for both players, different approaches to define the optimal solution paths are considered. Modifying the problem spaces and the solution paths could be extremely useful in order to make inferences about the strategy applied by the problem-solvers. An application of this procedure on tic-tac-toe is presented. Problem spaces for the game at issue were obtained by using the operation proposed by Simon and Newell (1972) and by Crowley and Siegler (1993). Then, a set of problems was obtained from these problem spaces, and different knowledge structures were derived.

Bayesian models as conditionally specified probabilistic structures. Highlights on the compatibility condition

Luigi Burigana

University of Padua

A probabilistic model is said to have conditional specification when the postulates shaping it are statements concerning conditional distributions and conditional independences among the random variables involved in the model, rather than statements concerning marginal distributions and marginal independences. Hierarchical Bayesian models, Markov random fields, and Bayesian networks make notable examples of that general concept. Given a set of postulates that specify conditional distributions within definite subsets of a general set of variables, there is no absolute assurance that those distributions are mutually compatible, that is, that there exists a consensus distribution (over the whole set of variables) from which all those conditional distributions may be deduced. This is known as the compatibility problem concerning conditionally specified probabilistic models. The problem is about the discovery of conditions that characterize mutual compatibility within given sets of conditional distributions and the construction of effective procedures for practically testing compatibility and gaining knowledge of consensus joint distributions, if these exist. During the past three decades, insightful research has been developed on these questions, especially within statistical science (a book by Arnold, Castillo, and Sarabia of 1999 is representative of the results obtained, up to that year). In my poster, I present definite results concerning the compatibility requirement as specifically related to some Bayesian models available in perceptual and cognitive psychology.

The Inter-Relation of Processing and Storage in Working Memory cannot be explained by Cognitive Load

Jan Göttmann[a] *& Gidon T. Frischkorn*[b]

[a] Heidelberg University, [b] University of Zurich

Current theories of working memory (WM) such as the Time-Based-Resource-Sharing-Model (Barrouillet & Camos, 2007) assume that the storage and processing (e.g. updating) of memory items in WM

are inter-related processes. Specifically, the TBRS-Model assumes that both maintenance of memory items and concurrent processing rely on the same attentional resource, which can only be utilized consecutively. This inter-relation is specified by cognitive load (CL), the ratio of specific task time t-alpha to total time T of a task. Thus, if CL is held constant, there should be no main effect of additional processing steps and no interactions between memory demands. Memory items shouldn‘t suffer from temporal decay, because there is sufficient freetime for attentional refreshing. To test this hypothesis, we conducted an experiment with N = 39 subjects who had to memorize 3 to 7 letters with concurrent working memory updating at a constant CL. We found decreasing accuracies (ACC) and increasing reaction times (RT) for additional processing steps and significant interactions on both measures. To validate the results, we estimated parameters of a Drift-Diffusion-Model (DDM). The most parsimonious model only varied the drift parameter v. According to the results on ACC and RT, there was a significant decrease of drift-rates v for additional processing steps and memory load. We conclude that CL suggested by the TBRS-Model does not describe the interaction between processing and storage sufficiently.

Testing process theories of causal reasoning using temporal uncertainty and response time models

Ivar Kolvoort & Leendert van Maanen

University of Amsterdam

Causal knowledge, i.e. knowledge about causal relationships, has been found to impact task performance ranging from reasoning and decision-making to categorization and learning. The idea that causality is crucial for the human cognitive system is gaining more traction and this is reflected by an increase in interest from cognitive scientists in the topic. In typical studies on causal reasoning, participants are provided with a network of a few variables that are causally related to another, after which they are asked to judge the probability of one variable having a value conditional on the other variables. Recent theories suggest that the cognitive processes involved in such causal judgments require sampling over possible states of the causal system, rather than exact computation of the (conditional) probabilities of certain outcomes. Judgments are then

based on the relative frequencies within the obtained samples. This sampling theory ascribes systematic deviations from normative models in human causal judgments to noise and bias in the sampling process. The sampling theory predicts participants to be able to provide an evidence-based probability estimate at any time during the sampling process, as judgments can be made based on any number of samples. Hence, uncertainty in the time that can be used for sampling should not affect the accuracy of judgments, only the available time should matter as this impacts the amount of samples that can actually be generated. However, if participants use explicit calculation of posterior probabilities or certain rule-based heuristics, temporal uncertainty affects the probability judgments because these computations would need to be finished before a judgment can be made. These processes require a specific amount of time, and hence could lead to guessing if this time was not available. In a new experiment, we rigorously tested these predictions by manipulating the amount of samples a participant could take to indicate the probability of a certain outcome in a causal network. Both a temporal deadline before which participants had to respond and uncertainty about this deadline were manipulated in a causal inference task. Sampling theory predicts that temporal uncertainty should not affect accuracy, since the number of samples only depends on the available time. In addition, if participants generate different numbers of samples due to time pressure, then changes in accuracy should be reflected in changes of the threshold parameter of a sequential sampling model, which represents the amount of evidence required to commit to a decision. This prediction was tested by fitting the spatially continuous diffusion model (SCDM) to the data. The SCDM was necessary because participants provided probability judgments on a continuous scale. With the current work, we aim to test predictions from verbal theories of causal reasoning, by introducing response times and response time models as an analytical tool in this domain.

Do achievement-motivated individuals increase their effort after receiving negative performance feedback? – A diffusion model analysis

Veronika Lerche, Julia Karl, Anne Buelow, Elena Melchner, Andreas Voss & Silke Hertel

Heidelberg University

In the field of motive research, the application of mathematical models is not yet common. Critically, if analyses of data are exclusively based on behavioral variables such as mean RT or accuracy rate, incorrect interpretations might result. One up-to-date uncontroversial assumption from the achievement motive literature is that individuals high in the implicit achievement motive invest more effort after they have been given negative, intraindividual performance feedback (e.g., Brunstein & Hoyer, 2002). This interpretation resulted from an analysis of mean RTs. We tried to replicate this finding in two studies (N_1 = 108, N_2 = 104). Thereby, we applied the diffusion model (Ratcliff, 1978) to investigate in which parameters achievement-motivated individuals differ from their less motivated counterparts. Interestingly, in both studies, the achievement-motivated individuals reduced their boundary separation from the first to the second part of the task. Thus, rather than investing more effort, they seem to have become less cautious. We dicuss this finding in the context of emotion regulation strategies.

Neuro-Fuzzy Inference System predicts heterogeneity of eye movements in psychiatric populations

Matteo Orsoni[a] , Federica Ambrosini[a] , Sara Garofalo[a] , Giovanni Piraccini[b] , Roberta Raggini[b] , Rosa Sant'Angelo[b] & Mariagrazia Benassi[a]

[a] University of Bologna, [b] Azienda Unità Sanitaria Locale della Romagna

Eye movements have been analyzed as an endophenotype of schizophrenia by applying inferential linear models. However, linear models failed to explain the heterogeneity of eye movements data in these populations. We aim to evaluate the performance of Neuro-Fuzzy Inference System modeling (ANFIS) in predicting pursuit eye

movements efficiency in psychiatric inpatients. Eighty-five psychiatric patients took part in the study (mean age 44.6 years, 54 males). Smooth pursuit eye movements were recorded with Eyetribe infrared system. The gain (ratio between eye movements velocity and stimulus velocity) was the measure of eye movements efficiency. Age, vision perception, symptoms severity and general cognitive functioning (evaluated with MMSE) were used as predictors. Five new subsamples were generated from the original data by bootstrapping resampling procedure for the testing phase of the model. Matlab's NeuroFuzzy Designer was used to verify the ANFIS performance. A learned model trained from original data showed high accuracy (RMSE=.04) and this result was confirmed also in the testing phase (RMSE=.05). Our findings show that ANFIS is a feasible tool to study how to predict eye movements efficiency in psychiatric populations. 3D plot surfaces have allowed to view the relationships between our variables of interest. Neuro-Fuzzy Inference System seems a possible solution to face the problem of non linear relations within complex phenomenon.

Fisherian and Bayesian approaches to eye movements analysis in psychotic disorders

Matteo Orsoni[a] , Sara Garofalo[a] , Federica Ambrosini[a] , Giovanni Piraccini[b] , Giovanni De Paoli[b] , Rosa Sant'Angelo[b] , Sara Giovagnoli[a], Roberto Bolzani[a] & Mariagrazia Benassi[a]

[a] University of Bologna, [b] Azienda Unità Sanitaria Locale della Romagna

Patients with schizophrenia and other psychotic disorders exhibit different types of eye movements (EM) deficits (Gooding & Basso, 2008; Reilly at al., 2014). However, large individual differences have been reported in literature, demonstrating that ocular motility in these patients is highly heterogeneous (Strakowski, et al., 2010). The reliability of EM studies in this domain is limited by the almost exclusive use of Fisherian statistics, an approach that while allowing to generalize a result from sample to whole population, reveal weakness in the analysis of population heterogeneity (Wagenmakers & Grunwal, 2007; Ziliak & McCloskey, 2008). More recent lines of research suggest the use of complementary statistical methods that take individual performance into account. On the one hand, Fisherian

statistics indicate if the differences or effects are statistically significant, and are based on the available data only. The Bayesian approach, on the other hand, can integrate this knowledge by informing about how strong the effect is, and by comparing the probability of different experimental hypotheses, given the obtained data. Furthermore, by estimating the probability of each possible measure, this latter approach can take into account the heterogeneity of performances. Although the theoretical complementarity of the two statistical approaches has been already discussed in literature, there is little empirical evidence for it. For this reason, the present study aimed to contrast and compare the statistical results arising from Fisherian and Bayesian approaches on the differences between EMs of a group of patients with schizophrenia spectrum and a healthy control group. The results showed that, while Fisherian statistics confirmed the presence of abnormal EMs in the patient group Bayesian statistics were better suited to clarify the extent of such difference and which type of EM is more strongly impaired. A deeper investigation of the role played by different EMs into schizophrenia and other psychotic disorders constitutes an interesting framework that may bring a relevant contribution to the current understanding of this spectrum of pathologies. These results may have implications in assessing, understanding and treating psychiatric disorders in the future.

Beyond one test: An R-Package for Item Pool Visualization

Nils Petras[a] , Michael Dantlgraber[b] & Ulf-Dietrich Reips[b]

[a] University of Mannheim, [b] University of Konstanz

To visually display tests on latent constructs with facets, Dantlgraber, Stieger, and Reips (2018) developed the concept of Item Pool Visualization (IPV). The radial IPV charts represent the content of tests, their facets, and related tests. IPV facilitates content based comparisons of different item sub-pools (e.g. similar facets of different tests or substantially different tests using similar names). It enables the understanding of factor structures in a new way, by introducing the center distance statistic. The center distance (cd) of an item (i) indicates how much more variance of that item is explained by a specific factor (s), extracted from a small set of items, compared to a more general factor (g) – that comprises additional

items: $cd_i = \lambda_{is}^2 / \lambda_{ig}^2 - 1$. The center distance can be used to compare the explanatory power of the latent construct, the general factor of each test, and the facets of the tests. In other words, how much better a more specific term describes the item, compared to a more general term. Additionally, latent correlations between tests and facets are displayed. IPV uses confirmatory structural equation models (SEMs). It is meant to inspire ideas on test content comparison, the meaning of test and facet labels, and the way items are located in existing tests and their facets. IPV can be used to review the test landscape in a similar way network diagrams are used. Network diagrams have been used to explore sets of items or symptoms by free restructuration. Latent structures are omitted in network diagrams and replaced by data-driven visual clustering. IPV also uses distance to structure item pools, but still reflects the way items are allocated to latent constructs in practice. For the assessment of the explanatory gain by facets over and above a general factor, bifactor models have been widely used as a general solution. While bifactor models are used to separate a general factor from specific factors in a single model, IPV compares the single factor SEM with the group factor SEM. Therefore, IPV is allowing for correlations between factors, and the estimated factor structure more closely resembles the common usage of sum scores in tests. Bifactor models force the dissection of explained item variance by the general factor and a specific factor. IPV compares the explanatory power of the general factor and the facet when estimated in their own right. In that sense, IPV can be a useful alternative approach to bifactor models. The current work implements IPV in R. The IPV package empowers the user to create each of the three radial IPV chart types by using a single, easy to understand function call. Additionally, the data input is facilitated by input functions. The package allows for detailed customization but simultaneously provides sensible, dynamic default values for all graphical parameters.
(https://github.com/NilsPetras/IPV)

Taming the Intractable: Generative Deep Learning for Universal Parameter Estimation

Stefan Radev, Ulf Mertens, Veronika Lerche & Andreas Voss

Heidelberg University

Mathematical models of cognition are formal descriptions of psychological theories allowing a clear and unambiguous way to formulate and test scientific hypotheses. Evidence accumulator models (EAMs) are a popular family of mathematical models about the cognitive processes underlying (perceptual) decision-making and response time generation. The parameters of many interesting EAMs, such as the Leaky Competing Accumulator (LCA), the Feed-Forward Inhibition (FFI), or Lévy-Flight-based models, are hard or even impossible to estimate from empirical data because of intractable likelihoods. We propose a novel, fast, and universally applicable method for likelihood-free Bayesian parameter estimation with a concrete focus on EAMs. The method relies on a deep generative neural network that enables us to approximate the full posterior over parameters by optimizing the posterior directly. Thus, the method can be regarded as inherently Bayesian. We present the numerous advantages of the method and apply it to both simulated and empirical response-time data to obtain state-of-the art parameter recovery. Further, we present a beta-version of a flexible, easy-to-use software with a graphical user interface, which enables researchers to use pre-trained DNNs on various existing EAMs or create and load their own models for non-standard experimental designs.

Randomized Responses within a Curtailed Sampling Plan: Efficient assessment of sensitive attributes

Fabiola Reiber & Rolf Ulrich

University of Tübingen

Randomized Response Techniques (RRTs) aim at increasing the validity of measuring sensitive attributes by eliciting more honest responses through anonymity protection of respondents. This anonymity protection is achieved by implementing randomization in the questioning procedure. On the other hand, this randomization increases the sampling variance of the resulting estimates and,

therefore, increases sample size requirements. The present work aims at countering this drawback by combining RRTs with curtailed sampling, a sequential sampling design in which sampling is terminated as soon as sufficient evidence to decide on a hypothesis is collected. In contrast to open sequential designs, the curtailed sampling plan includes the definition of a maximum sample size. We show that resources can be saved using this approach and discuss the advantages and disadvantages in light of differing research foci.

Enhancing lie detection with sequential sampling models

Lars M. Reich, Bartosz Gula, Gáspár Lukács, Deniz Tuzsus & Rainer W. Alexandrowicz

University Klagenfurt

Previous research has shown that sequential sampling models capture well the Speed Accuracy Tradeoff (SAT) in various speeded decision tasks. In the current study the aim was to analyze, whether the Drift Diffusion Model (DDM) and the Linear Ballistic Accumulator Model (LBA) can be applied to other speeded decision tasks in the context of lie detection. In a Reaction Time based Concealed Information Test (RT-CIT) participants were randomly assigned two groups (guilty/innocent). The guilty group had to commit a mock crime while the innocent group was provided with an instruction to copy a sentence from a book in the library. All participants (N=90), regardless of the real task they had do perform, were instructed to report the same story, namely having copied a sentence from a book in the library. The guilty group was split into two subgroups, which were trained on how to fake the RT-CIT in two different ways (speed up on probe items/slow down on irrelevant items). Data analysis involved investigating the fit of LBA and DDM to present data before and after fake training. Further analysis attempts to successfully distinguish between the density distribution of the innocent and the guilty participants.

The automatic model of dynamic characteristics of the person

Tatiana Savchenko & Galina Golovina

Russian Academy of Sciences, Moscow

To simulate the agent we used Krylov's Automat, who has q states (the memory volume). The transition from state to state depends on the gain or loss and number of the previous step. This automat is asymptotically optimal in a stochastic environment Markov, given by the probability of winning (P1, ..., Pm). This means that the marginal expectation of winning is equal to the maximum probability automaton Pi in the medium fi i = 1, ..., m Lim M = max Pi. In modeling behavior of the agent, the automat introduced quantity q. It is called the memory depth machine. The meaning of this parameter is as follows. The larger q, the more inertial machine, for the greater consistency of losses forced a person to change the action. Intuitively, the greater the inertia of the automat, the closer it is to ensure that by choosing the best action in this environment, it continues to perform only for him.As the depth increases the feasibility of memory, behavior of an automaton in stationary environments. Conversely, at low values of q over the operation of the machine is exposed to losses that could translate into automatic new action.The probability of winning in the environment (stochastic chain Markov) transition probabilities agent from one state to another, as well as the depth of memory (for the simulation of each type) were found in the empirical study using the frequency characteristics of the types.The probability of agent transitioning from one state to another is determined by empirical research. That is this frequency (probability) of types such as social arranger and self-acceptance.A marginal gain in cases of the modeling of dynamic types of satisfaction with life is a life satisfaction assessment of this type.Satisfaction with each type of agent is equal to the amount of winnings in each action. Overall, life satisfaction is calculated as the amount of winnings of agents of all types.It was proposed by a theoretical model of micro dynamic SQL types. Social arranger Social hosted by the random environment is defined as hosted – in a random environment with a high probability of winning p1, and disorder - a random medium with a low probability of winning p2.Satisfaction with the life agent is defined by the depth of the memory machine: the state of satisfaction corresponds to q = 1 corresponds to the condition of maximum satisfaction of q = 10.The

probability of a change agent environment is defined by the values r1 and r2, with values changing from β and α 0.1 to 0.9, and finally, the probability for another round of memory (above and below).Thus, there are six sets of probabilities for each type, each of which consists of seven settings - p1, p2, q, β, α, r1, r2.

Bayesian inference for a model of eye-movement control during reading

Stefan Seelig, Ralf Engbert, Sebastian Reich & Maximilian Rabe

University of Potsdam

The processes involved in the generation of eye-movements and written language processing (reading) are widely assumed to dynamically interact at several levels. Although several computational models of eye-movements during reading exist, the methods used for parameter estimation are often weak as a result of model complexity. The fact that underlying data are time-ordered fixation-sequences is typically ignored. Here we present the construction of a likelihood function for a Bayesian framework of parameter estimation. We apply this approach to experimental datasets and successfully model interindividual differences in gaze behavior using the SWIFT model of saccade generation. As the mathematical model structure prohibits a closed-form expression of the likelihood function, we propose a combined pseudo-marginal likelihood with approximate Bayesian computation to yield viable and computationally efficient likelihood computations. We then use advanced Markov Chain Monte Carlo procedures (Vihola, 2012; Vrugt, 2009) to recover and estimate parameters on a participant level for simulated and experimental data respectively. By comparing experimental data and simulations based on previously estimated parameters we can demonstrate the capability of this approach in modeling intricate interindividual differences, as well as its value for model building and evaluation in general.

References:

Vihola. M. (2012). Robust adaptive Metropolis algorithm with coerced acceptance rate. *Statistics and Computing, 22*, 997-1008.

Vrugt, A., Ter Braak, C.J.F., Diks, C.G.H., Robinson, B., Hyman, J. & Higdon, D. (2009). Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences & Numerical Simulation, 10*(3), 273-290.

Conjoint structures of polytomous items and the empirical requirements of additivity

Luca Stefanutti

University of Padua

Adding item responses to form “scores” is a very pervasive and popular practice for analyzing data obtained through psychological tests and questionnaires. However, empirical evidence supporting the assumption that some form of concatenation exists among item responses is rarely provided. Many attempts exist in the form of the various parametric or nonparametric item response theory extensions of the Rasch model. All of them adhere to the “ability minus difficulty” view of additivity in which the latent trait is assumed to be continuous. Unfortunately, those models have no convincing empirical methods for demonstrating neither additivity nor continuity (see, e.g., Heene, 2013; Michell, 2008). This talk is on the empirical requirements that must be satisfied for concluding that some additive numerical representation exists (and is unique) for the responses of a finite set of polytomous items. In the simplest case of two items (two components) and a rating scale with a finite set of totally ordered levels, such requirements are already established and can be found in the first volume of the Foundations of Measurement (Krantz, Luce, Suppes, Tversky, 1971). Following a conjoint measurement approach, the case of n items and k ordered levels is examined and a special axiom is introduced, which is found to be sufficient for having restricted solvability in the finite case. The introduced axiom can be viewed as a generalization of the “equally spaced intervals” assumption of the two-component case to the n-components one. It induces an equivalence relation (named the “indifference relation”) on the set of the response patterns, which is only preserved by positive linear transformations of the scale. The necessary requirement for concluding additivity of item responses is to provide empirical evidence of the existence of such an indifference relation.

Otherwise, any monotone increasing transformation goes. However, nonlinear increasing transformations have a critical impact on the results of statistical inference applied to scores. For instance, parametric and nonparametric statistical tests of the difference between population means, for how robust or powerful they can be, may fail anyway. This is because, in general, monotone scale transformations do not preserve the difference between population means. A series of simulation studies show that nonlinear monotone scale transformations do not preserve the indifference relation on the set of response patterns, with the consequence that the size of the difference between population means strongly depends on the applied transformation.

Slower, but why? A meta-analysis on age differences in diffusion model parameters

Maximilian Theisen, Veronika Lerche, Mischa v. Krause & Andreas Voss

Heidelberg University

Older adults typically show slower response times in basic cognitive tasks than younger adults. The diffusion model allows to estimate cognitive components of decision making and can be used to clarify why older adults may react more slowly. Main components of the diffusion model are the speed of information uptake (*drift rate*), the degree of conservatism regarding the decision criterion (*boundary separation*), and the time taken up by non-decisional processes (i. e. encoding and motoric response execution; *nondecision time*). While the literature shows consistent results regarding higher boundary separation and longer nondecision time in older adults than younger adults, the results are more complex when it comes to age differences in the drift rate. We conducted a multi-level meta-analysis to identify possible sources of this variance. As possible moderators, we included task difficulty and task type. We found that age differences in drift rate differ by task type and task difficulty. Older adults were inferior in drift rate in perceptual and memory tasks, but they performed better than younger adults in lexical decision tasks. Older individuals benefitted if task difficulty was higher, however only in the perceptual and lexical decision tasks. In the memory tasks, task

difficulty did not moderate the effect of age. The finding of higher boundary separation and longer nondecision time in older adults than younger adults generalized over task type and task difficulty. The results of our meta-analysis are consistent with recent findings of a more pronounced age-related decline in memory than in vocabulary performance.

Using the diffusion model to assess dark personality

Mischa v. Krause & Andreas Voss

Heidelberg University

In recent years, there has been an increase of interest in so so-called dark personality traits, ie. traits that manifest in socially undesirable or even downright malevolent behavior. Such traits were typically assessed using self-report questionnaires, with the most popular instruments trying to assess the Dark Triad of psychopathy, narcissism and Machiavellism. While these instruments have been fundamental in advancing the study of dark personality, they share the problems inherent in all self-report measures, for example the reliance on conscious introspection and easy fakeability. These issues seem especially important given the fact that the traits assessed are by their very definition socially undesirable. We introduce a new instrument based on simple binary decisions under time pressure - does this adjective describe me well? Ratcliff's diffusion model is employed in order to achieve - in the model parameter drift rate - a more pure measure of speed of information uptake in these decisions than simple RTs. The difference in drift rates for "dark" and "light" adjectives is used as an estimate of dark personality. We present initial data that points towards concurrent, incremental and predictive validity of the measures obtained.

List of Authors

Program

Monday, August 5

08:30-10:10	**Judgement & Decision Making** Chair: *Henrik Olsson* Dynamic Choice with Status Quo: Theory and Design of Efficient Experiment *Hassan Nosratabadi & Francois Maniquet* Quantum rotation: a new method for capturing a change of perspective *Thomas Hancock & Stephane Hess* The disjunction effect in two-stage gamble experiments *Jan Broekaert, Jerome Busemeyer & Emmanuel Pothos* A Hidden Markov Model approach to model Mouse-Tracking Data *Marco D'Alessandro, Luigi Lombardi & Antonio Calcagnì* The Strategy Aggregation Effect in Group Judgment *Henrik Olsson & Mirta Galesic*
10:10-10:40	Coffee Break
10:40-12:20	**Bayesian Statistics** Chair: *Martin Lages* Methodological flexibility in prior elicitation and its effects on Bayesian model comparison *Angelika Stefan* Sample Size Determination for the Bayesian t-test *Qianrao Fu, Herbert Hoijtink & Mirjam Moerbeek* Bayesian Evaluation of Informative Hypotheses *Herbert Hoijtink* Bayesian analysis of information used during decision making *Tillmann Nett, Nadine Nett & Andreas Glöckner* Variance Constraints For Hierarchical Signal Detection Models *Martin Lages*
12:20-13:40	Lunch Break
13:40-14:40	**Keynote** Qualitative vs. Quantitative Individual Differences: Implications for Cognitive Control *Jeff Rouder*

14:40-15:10	Coffee Break with Sports
15:10-16:30	**Knowledge Space & Learning I** Chair: *Jean-Paul Doignon* Empirical Distinctness of Skill Map Based Knowledge Structures *Andrea Spoto & Luca Stefanutti* Extracting Partially Ordered Clusters from ordinal polytomous data: A comparison between *k*-modes and *k*-median algorithms *Debora de Chiusole, Andrea Spoto & Luca Stefanutti* Extracting quasi-orders from polytomous data by a minimum discrepancy approach *Luca Stefanutti* From Italian menus to resolutions of learning spaces *Jean-Paul Doignon*
16:30-17:30	**Poster Session with Refreshments & Snacks** A Sequential Sampling Model for Visual Search RT Distributions *Steven Blurton* A problem space approach to studying strategic planning in board games *Andrea Brancaccio & Luca Stefanutti* Bayesian models as conditionally specified probabilistic structures. Highlights on the compatibility condition *Luigi Burigana* The Inter-Relation of Processing and Storage in Working Memory cannot be explained by Cognitive Load *Jan Göttmann & Gidon T. Frischkorn* Testing process theories of causal reasoning using temporal uncertainty and response time models *Ivar Kolvoort & Leendert van Maanen* Do achievement-motivated individuals increase their effort after receiving negative performance feedback? – A diffusion model analysis *Veronika Lerche, Julia Karl, Anne Buelow, Elena Melchner, Andreas Voss & Silke Hertel* Neuro-Fuzzy Inference System predicts heterogeneity of eye movements in psychiatric populations *Matteo Orsoni, Federica Ambrosini, Sara Garofalo, Giovanni Piraccini, Roberta Raggini , Rosa Sant'Angelo & Mariagrazia Benassi*

	Fisherian and Bayesian approaches to eye movements analysis in psychotic disorders *Matteo Orsoni, Sara Garofalo , Federica Ambrosini, Giovanni Piraccini , Giovanni De Paoli, Rosa Sant'Angelo, Sara Giovagnoli & Roberto Bolzani* Beyond one test: An R-Package for Item Pool Visualization *Nils Petras, Michael Dantlgraber & Ulf-Dietrich Reips* Taming the Intractable: Generative Deep Learning for Universal Parameter Estimation *Stefan Radev, Ulf Mertens, Veronika Lerche & Andreas Voss* Randomized Responses within a Curtailed Sampling Plan: Efficient assessment of sensitive attributes *Fabiola Reiber & Rolf Ulrich* Enhancing lie detection with sequential sampling models *Lars M. Reich, Bartosz Gula, Gáspár Lukács, Deniz Tuzsus & Rainer W. Alexandrowicz* The automatic model of dynamic characteristics of the person *Tatiana Savchenko & Galina Golovina* Bayesian inference for a model of eye-movement control during reading *Stefan Seelig, Ralf Engbert, Sebastian Reich & Maximilian Rabe* Conjoint structures of polytomous items and the empirical requirements of additivity *Luca Stefanutti* Slower, but why? A meta-analysis on age differences in diffusion model parameters *Maximilian Theisen, Veronika Lerche, Mischa v. Krause & Andreas Voss* Using the diffusion model to assess dark personality *Mischa v. Krause & Andreas Voss*
17:40-19:30	Philosopher's Walk
19:30-22:00	Dinner at restaurant *PalmBräu Gasse* (self-pay; address: Hauptstraße 185)

Tuesday, August 6

08:30-10:10	**Response Times I – Diffusion Modeling** Chair: *Andreas Voss* Thinking fast, not slow: A drift diffusion model account of belief bias *Anna-Lena Schubert, Mário B. Ferreira, André Mata & Ben Riemenschneider* A comparison of conflict diffusion models in the flanker task through pseudo-likelihood Bayes factors *Nathan Evans & Mathieu Servant* Diffusion models with time-dependent drift rates: A partial differential equation solution (Part 1) *Markus Janczyk, Rolf Ulrich & Thomas Richter* Diffusion models with time-dependent drift rates: Numerical accuracy and efficiency in simulation and parameter estimation (Part 2) *Thomas Richter, Markus Janczyk & Rolf Ulrich* A Lévy-Flight Model of Decision Making *Andreas Voss & Marie Wieschen*
10:10-10:40	Coffee Break
10:40-12:20	**Cognitive & Statistical Models** Chair: *Nadia Said* Causal Interpretation of Statistical Models - Why we shouldn't ignore the scientific philosophers *Andreas Klein* Factorial invariance and orthogonal rotation *Silvia Testa & Luca Ricolfi* Traps and tricks in Monte-Carlo simulation-based parameter estimation of advanced mathematical models *Yiqi Li, Martin Schlather & Edgar Erdfelder* The analysis of the response profile of Motion and Form coherence tests by means of half normal psychophysical function *Sara Giovagnoli, Roberto Bolzani, Luca Mandolesi, Kerstin Hellgren, Sara Garofalo & Mariagrazia Benassi* Understanding Belief Polarization - An Agent-Based Modeling Approach *Nadia Said, Debora Fieberg, Helen Fischer, Andreas Potschka & Christian Kirches*

12:20-13:40	Lunch Break
13:40-14:40	**Keynote** A Statistician's Botanical Garden - The Ideas behind Decision Trees and Random Forests *Carolin Strobl*
14:40-15:10	Coffee Break with Sports
15:10-16:30	**Knowledge Space & Learning II** Chair: *Jürgen Heller* Learning to Compare *Louis Narens* Unique skills assessment via minimal competence models *Pasquale Anselmi, Jürgen Heller, Luca Stefanutti & Egidio Robusto* A graphical taxonomy of assessment models *Stefano Noventa & Jürgen Heller* Representing misconceptions within polytomous Knowledge Structure Theory *Jürgen Heller*
16:30-17:00	Coffee Break
17:00-18:20	**Modelling Choices** Chair: *Lena Wollschlaeger* Introduction of right action on extensive structures for intertemporal choice *Yutaka Matsushita* How to deal with rational intransitive choices *Reinhard Suck* Axiomatic properties of bad decision *Kazuhisa Takemura* A computational cognitive process model for best-worst choice *Lena Wollschlaeger & Adele Diederich*
18:20-18:50	Business Meeting
20:00-23:00	Conference Dinner at BräuStadl (included in registration fees; address: Berliner Straße 41)

Wednesday, August 7

08:30-10:10	**Response Times II** Chair: *Matthias Gondan* Evidence accumulation in same-different judgments *Andrew Hendrickson, Danielle Navarro & Chris Donkin* The Dynamics of Decision Making During Goal Pursuit *Timothy Ballard, Andrew Neal, Simon Farrell & Andrew Heathcote* Testing the k-modal race model inequality *Luigi Lombardi & Hans Colonius* Generalizing the Memory Measurement Model to n-AFC recognition retrievals *Gidon T. Frischkorn & Klaus Oberauer* Incorrect responses in the response time interaction contrast *Matthias Gondan*
10:10-10:40	Coffee Break
10:40-12:20	**Psychometrics & Psychophysics** Chair: *Julia M. Haaf* Fuzzy Item Ambiguity Analysis in psychological testing and Measurement *Hojjatollah Farahani & Parviz Azadfallah* Towards meaningful inferences from attitudinal thermometer ratings *Yung-Fong Hsu, Michel Regenwetter & James Kuklinski* Alternatives to the Inverse Wishart distribution in Bayesian hierarchical IRT models *Christoph Koenig & Alexander Naumann* An updated concept of evidence based on Bayes' law, which explains decision making for sensory tasks with numerous complicated objects *Valentin Shendyapin & Irina Skotnikova* Do Items Order? The Psychology of IRT Models *Julia M. Haaf*
12:20-13:40	Lunch Break

13:40-14:40	**Keynote** A computational cognitive modeling approach to understanding and reducing errors in medical image-based decision-making *Jennifer Trueblood*
14:40-15:10	Coffee Break with Sports
15:10-16:10	**Multinomial Models** Chair: *Daniel W. Heck* Efficient Hypothesis Tests in Multinomial Processing Tree Models: A Sequential Probability Ratio Test for the Randomized Response Technique *Martin Schnuerch & Edgar Erdfelder* Response Time Extended Multinomial Processing Trees in R *Raphael Hartmann, Karl Christoph Klauer & Lea Johannsen* Bayesian Inference for Multinomial Models with Linear Inequality Constraints *Daniel W. Heck & Clintin P. Davis-Stober*

	Monday	Tuesday	Wednesday
08:30-10:10	Judgement & Decision Making	Response Times I	Response Times II
10:10-10:40	Coffee Break	Coffee Break	Coffee Break
10:40-12:20	Bayesian Statistics	Cognitive and Statistical Models	Psychometrics & Psychophysics
12:20-13:40	Lunch Break	Lunch Break	Lunch Break
13:40-14:40	Keynote Jeff Rouder	Keynote Carolin Strobl	Keynote Jennifer Trueblood
14:40-15:10	Coffee Break with Sports	Coffee Break with Sports	Coffee Break with Sports
15:10-16:30	Knowledge Space & Learning I	Knowledge Space & Learning II	15:10-16:10 Multinomial Models
16:30-17:00	16:30-17:30 Poster Session with Refreshments & Snacks	Coffee Break	
17:00-18:20		Modelling Choices	
Evening Event	17:40-19:30 Philosopher's Walk 19:30-22:00 Dinner at *PalmBräu Gasse*	18:20-18:50 Business Meeting 20:00-23:00 Conference Dinner at *BräuStadl*	