

Online SMiP-IOPS Conference, 8th July, 2021

09:00 – 09:15: Welcome and Introduction

09:15 – 10:45: Talk Session 1, Chair: Rob Meijer

Lodder, Paul (IOPS) *The psychometrics of type D personality*

Petras, Nils (SMiP) *Measuring the domain-specific content of psychological constructs*

Meijerink, Marlyne (IOPS) *Dynamic relational event modeling: Testing, exploring, and applying*

10:45 – 11:00: Short Break

11:00 – 12:30: Talk Session 2, Chair: Andrea Kiesel

Stoevenbelt, Andrea (IOPS) *Time limits as potential source of gender bias in experimental settings:
The case of stereotype threat*

Laukenmann, Ruben (SMiP) *Process models of the weapon identification task*

Stump, Annika (SMiP) *Is it all about the feeling? Affective and (meta-)cognitive mechanisms
underlying the truth effect*

12:30 – 13:30: Lunch Break *

13:30 – 15:00: Talk Session 3, Chair: Dylan Molenaar

Stefan, Angelika (IOPS) *Bayes factor forecasts for continuous research design evaluation*

Schreiner, Marcel (SMiP) *Modeling binding effects in episodic memory: A comparison of five
approaches*

van den Akker, Olmo (IOPS) *Selective hypothesis reporting in psychology*

15:00 – 15:05: Group photo

15:05 – 15:20: Short Break *

15:20 – 16:30: Poster Session **

Presenters: *see attachment.*

16:30 – 16:45: Short Break *

16:45 – 17:45: Keynote Talk

Jeffrey N. Rouder *Model specification in everyday, run-of-the-mill statistical testing: Why it
is an open question; Why it matters; Why it is your problem*

17:45: End of official Program *

19:30: Game Night ***

Online SMiP-IOPS Conference, 9th July, 2021

09:00 – 10:30: Talk Session 4, Chair: Tanja Lischetzke

- | | |
|------------------------|---|
| Ernst, Anja (IOPS) | <i>Dynamic clustering: Classifying people through ecological momentary assessment</i> |
| Schmitt, Marcel (SMiP) | <i>Modeling emotion differentiation by means of Latent Markov Factor Analysis</i> |
| Haslbeck, Jonas (IOPS) | <i>Modeling psychopathology: From data models to formal theories</i> |

10:30 – 11:00: Coffee Break *

11:00 – 12:30: Talk Session 5, Chair: Beatrice Kuhlmann

- | | |
|-------------------------------|--|
| Park, Soogeun (IOPS) | <i>discovR: Classification model for data from multiple sources</i> |
| Pratiwi, Bunga Citra (IOPS) | <i>Predictive performance of psychological tests: Is it better to use items than subscales?</i> |
| Quevedo Pütter, Julian (SMiP) | <i>Different, but inextricable? An MPT modeling approach to diversion and similarity interference in episodic memory</i> |

12:30 – 13:30: Lunch Break *

13:30 – 14:45: Board Meeting IOPS (from 14:15 with SMiP representatives) /
Informal Meeting of PhD Candidates ***

14:45 – 15:15: Awards **** (Poster awards, presentation awards, paper awards)

15:15: End of Program *

Posters

| | |
|------------------------------|---|
| Constantin, Mihai (IOPS) | <i>A general Monte Carlo method for sample size analysis in the context of network models</i> |
| Fang, Qixiang (IOPS) | <i>Natural language processing for survey researchers: Can sentence embedding techniques improve prediction modelling of survey responses and survey question design?</i> |
| Fenn, Julius (SMiP) | <i>Exploring cognitive affective maps as a new mode of data collection</i> |
| Hoekstra, Ria (IOPS) | <i>Heterogeneity: Reality or illusion?</i> |
| Karimova, Diana (IOPS) | <i>Separating the wheat from the chaff: Bayesian regularization in dynamic social networks</i> |
| Liu, Xiaotong (SMiP) | <i>Testing sampling-based accounts of probability judgements using a ranking task</i> |
| Merhof, Viola (SMiP) | <i>Modeling dynamic response style effects</i> |
| Muradchanian, Jasmine (IOPS) | <i>How best to quantify replication success? A simulation study on the comparison of replication success metrics</i> |
| Rebholz, Tobias (SMiP) | <i>Adaptive advice taking: Mixed-effects weights and Bayesian stopping</i> |
| Scholz, David (SMiP) | <i>Beyond (low) agreeableness: A closer view on antagonistic psychopathology</i> |
| Schumacher, Lukas (SMiP) | <i>Understanding learning during multi-alternative decision-making</i> |
| Smith, Parker (SMiP) | <i>Motivation for a nonlinear extension of the diffusion model of conflict</i> |

Abstracts (Talks)

Ernst, Anja Franziska (IOPS)

“Dynamic clustering: Classifying people through ecological momentary assessment”

Studying within-person dynamics through time series models is becoming increasingly popular in the social sciences. Often, researchers are interested in summarizing the dynamics of several individuals into a common time series model. As dynamics can be rather heterogeneous across individuals, one needs sophisticated tools to express the essential similarities and differences across individuals. A way to proceed is to identify subgroups of individuals who are characterized by distinct differences in their dynamics. In my talk I will present two dynamic clustering models which aim to identify such latent subgroups. (1) The latent class vector-autoregressive model, which comes with an adaptive estimation procedure that allows different clusters to be modeled with qualitatively different time-series models. (2) The mixture multilevel vector-autoregressive model, which models the continuous between-person differences within the clusters. I will illustrate both models through concrete examples on empirical data sets.

Student Discussants: Bunga Citra Pratiwi and David Izydorczyk

Haslbeck, Jonas (IOPS)

“Modeling Psychopathology: From Data Models to Formal Theories”

There has been a surge of empirical research using network models to investigate mental disorders as complex systems. In this talk, I discuss how to best make use of those data models in order to construct theories that allow us to explain, predict, and control mental disorders. Specifically, I consider three ways in which one can use empirical findings (i.e., data models) to construct formal theories: (a) using data models themselves as formal theories, (b) using data models to infer formal theories, and (c) comparing empirical data models to theory-implied data models in order to evaluate and refine an existing formal theory. I argue that the third approach is the most promising path forward, and propose a framework to construct formal theories that provides a clear way forward for using empirical research to inform the generation, development, and testing of formal theories both in the domain of psychopathology and in the broader field of psychological science.

Student Discussants: Olmo van den Akker and Kilian Hasselhorn

Laukenmann, Ruben (SMiP)

“Process models of the weapon identification task”

The Weapon Identification Task (WIT) is a sequential priming paradigm, which assesses the effect of racial priming on visual discrimination between weapons (e.g., a gun) and innocuous objects (e.g., a tool). The Process Dissociation Procedure (PDP) is a widely used measurement model to estimate the influence of controlled and automatic cognitive processes on task performance. Klauer and Voss (2008) identified four process models based on the PDP that differ in their assumptions on the nature and interplay of these cognitive processes leading to the racial bias effect. We compared these process models using hierarchical response time extended Multinomial Processing Tree (MPT-RT) modeling. MPT-RT modeling shows a clear preference for two models – the Default Interventionist Model (DIM) and the Preemptive Conflict Resolution Model (PCRM) – which posit fast automatic and slow controlled process routes. Additional comparison of extended versions of the DIM and PCRM indicate a preference for the DIM postulating an early interference of automatic stereotype associations in weapon identification.

Student Discussants: Angelika Stefan and Fabiola Reiber

Lodder, Paul (IOPS)

“The Psychometrics of Type D personality”

The construct Type D personality is considered an important risk factor for adverse events in cardiovascular disease patients. Type D personality is conceptualized as high scores on the two personality traits negative affectivity (NA) and social inhibition (SI). Each trait is measured with seven items on a 0-4 Likert scale and Type D personality is commonly operationalized as scoring above a predetermined cutoff on both the NA and SI sum scores. This dichotomous approach has been criticized in the literature for not only resulting in less power but also for risking spurious associations with other variables. A less biased continuous approach operationalizes Type D personality as an interaction between the continuous NA and SI scores. This presentation has four aims. First, I present the results of simulation studies showing when and why the dichotomous approach is more biased than the continuous approach. Second, I discuss the impact of the biased dichotomous approach on the hundreds of published studies in the Type D literature using it. Third, I present the results of an individual patient data meta-analysis that reanalyzes 18 earlier published studies using the continuous approach. Lastly, I argue for modeling Type D personality as an interaction between the latent NA and SI variables using a structural equation model, taking into account the measurement error and skewness in the ordinal NA and SI item scores.

Student Discussants: Soogeun Park and Thomas Verliefde

Meijerink, Marlyne (IOPS)

“Dynamic relational event modeling: Testing, exploring, and applying”

The relational event model (REM) facilitates the study of network evolution in relational event history data, i.e., time-ordered sequences of social interactions. In real-life social networks, however, it is likely that network effects, i.e., the parameters that quantify the relative importance of drivers of these social interaction sequences, change over time. In these networks, the basic REM is not appropriate to understand what drives network evolution. Therefore, this research extends the REM framework with approaches for testing for and exploring of time-varying network effects. Firstly, we develop a Bayesian approach to test whether network effects change during the study period or not. A simulation study was conducted that illustrates that the Bayesian test accurately quantifies the evidence between a basic (‘static’) REM or a dynamic REM. Secondly, in the case of the latter, time-varying network effects can be studied by means of a moving window that slides over the relational event history. A simulation study was conducted that illustrates that the accuracy and precision of the estimates depend on the window width: narrower windows result in greater accuracy at the cost of lower precision. Thirdly, one challenge of the moving window REM is to determine the window width that can best capture the temporal effect dynamics. Therefore, we develop a Bayesian approach for determining window widths using the empirical network data. A simulation study was conducted that illustrates that estimation with empirically determined window widths achieves both good accuracy for time intervals with important changes and good precision for time intervals with hardly any changes in the effects. Finally, in an empirical application, it is demonstrated how the approaches in this research can be used to test for and explore time-varying network effects of face-to-face contacts at the workplace.

Student Discussants: Soogeun Park and Maike Arnold

Park, Soogeun (IOPS)

“discovR: classification model for data from multiple sources”

Having large sets of predictor variables from multiple sources concerning the same observation units and the same outcome variable is becoming increasingly common in behavioural research. Constructing a classification model from such data entails multiple objectives: classification of the outcome variable, variable selection and identification of processes at play underneath the predictors. These processes are of particular interest in the setting of multi-block data because they can either be associated individually with single data blocks or jointly with multiple blocks. Many methods have addressed the classification problem in high-dimensionality for a single block of data. However, the additional challenge of distinguishing the underlying processes from multi-block data has not received sufficient attention. To this end, we propose the method of discovR (distinctive and common sparse covariates regression). The method extends principal covariates regression to accommodate multi-block data and combines with generalized linear modeling framework to allow classification of a categorical outcome. In a simulation study, discovR resulted in outperformance compared to related methods commonly used in behavioural sciences.

Student Discussants: Esther Maassen and Monika Vaheoja

Petras, Nils (SMiP)

“Measuring the domain-specific content of psychological constructs”

Each psychological measure is a unique mixture of construct-relevant content. To disentangle a common core from specific content (domains), bi-factor models were developed. Problematically, “collapsing” domain factors, with little to no variance, are abundant in the bi-factor literature. I argue that there are good reasons to expect weak domain factors in practice, and that considering statistical power and parameter recovery is the key to successful research design. The current simulation study provides a systematic overview of statistical power and parameter recovery in bi-factor models. The results indicate, that reconsidering the design of the measure is often more appropriate than increasing sample size. Furthermore, the simulation clearly shows that uninterpretable results (non-convergence, negative trait variance estimates) coincide with tiny effect sizes. Moving forward, it will be discussed how bi-factor models could be used as comprehensive models across several measures. This locates the measures in a framework of the construct – and helps identify content for which power and recovery are problematic in individual measures.

Student Discussants: Dandan Tang and Angelika Stefan

Pratiwi, Bunga Citra (IOPS)

“Predictive Performance of Psychological Tests: Is it better to use Items than Subscales?”

Many studies and applications use psychological tests to predict various outcomes. This requires the researcher to generate a prediction rule from a psychological test. For multidimensional tests, the standard approach is to use the subscale scores in a multiple linear regression model estimated using ordinary least squares. Recently, instead of using subscale scores, several studies used separate item scores in combination with statistical learning methods to optimize predictive performance of these tests (Putka et al., 2018; Seeboth & Möttus, 2018). However, it is unclear whether this approach is always beneficial. The goal of this study is to identify in which situation(s) is it better to take into account the predefined structure in the items by forming subscale scores and then develop a prediction rule, in which situation(s) is it better to use the separate items, and in which situation(s) is it better to let the data decide on which scores to use in a prediction rule. We used several statistical methods to derive the prediction rules; ordinary least squares, elastic net, and component analysis (supervised principal components and principal covariates regression). We analyzed data from two empirical studies into the predictive validity of: 1) the Cognitive Emotion Regulation Questionnaire (CERQ) to predict depression scores (Garnefski et al., 2001), and 2) a personality inventory to predict job performance using reviewer ratings (Cubiks, 2018). Furthermore, a simulation study was also performed. Overall, results showed that it is preferable to use subscales rather than items to develop a prediction rule in situations such as when component scores are predictive and sample size is small.

Student Discussants: Marvin Neumann and Olmo van den Akker

Quevedo Pütter, Julian (SMiP)

“Different, but inextricable? An MPT modeling approach to Diversion and Similarity Interference in episodic memory”

Retroactive interference is assumed to play a major role in everyday forgetting. According to an influential model by Dewar et al. (2007), two different types of retroactive interference need to be distinguished: Diversion and Similarity Interference. Whereas Diversion Interference is thought to inhibit consolidation processes, Similarity Interference is assumed to result in a retrieval deficit. However, these crucial assumptions have not yet been tested empirically, most likely because consolidation and retrieval processes are not directly observable in behavioral data. Against this backdrop, I propose a multinomial processing tree (MPT) modeling approach to finally disentangle Similarity and Diversion Interference. After some theoretical and methodological considerations, I will present the results of a first online experiment and discuss future steps.

Student Discussants: Xynthia Kavelaars and Nikoletta Symeonidou

Schmitt, Marcel (SMiP)

“Modeling emotion differentiation by means of Latent Markov Factor Analysis”

Emotion differentiation (ED) is defined as an individual's tendency to describe his or her emotional experiences with specificity. To investigate ED, researchers typically use data from experience sampling studies in which participants repeatedly rate the intensity of multiple emotions. In the present talk, I would like to present my dissertation project aimed at testing an alternative modeling approach to investigate within-person and between-person variability in ED by applying latent Markov factor analysis (LMFA; Vogelsmeier et al., 2019). After shortly discussing some drawbacks of methodological approaches to operationalize ED taken in previous research, I will discuss the statistical framework of LMFA and its applicability to modeling ED. Finally, I will provide an insight into a currently running ambulatory assessment study, the data of which I will use to validate LMFA for modeling ED.

Student Discussants: Anja Ermst and Gloria Grommisch

Schreiner, Marcel (SMiP)

“Modeling Binding Effects in Episodic Memory: A Comparison of Five Approaches”

Experienced events consist of several elements that need to be bound together to represent the event in a coherent manner. Such binding process lead to a stochastic dependency of the retrieval of event elements. Several approaches for modeling these dependencies have been proposed. In a simulation study, the contingency-based approach by Horner and Burgess (2013), two related approaches using Yule's Q, a newly proposed IRT-based approach using the Q3 statistic (Schreiner et al., 2021), and a variant of the IRT-based approach using nonparametric estimation of Q3 (Debelak & Koller, 2020) are compared regarding their empirical detection rates of dependencies and differences in dependency and their susceptibility to different levels of overall memory performance. The different approaches are also applied to an empirical dataset to evaluate the congruence of inferences drawn from empirical data.

Student Discussants: Qixiang Fang and Monika Wiegelmann

Stefan, Angelika (IOPS)

“Bayes Factor Forecasts for Continuous Research Design Evaluation”

Well thought-out study designs are at the heart of experimental research. They can ensure that studies strike a balance between efficiency and informativeness, and that study conclusions are logically sound. However, researchers often need to plan studies based on insufficient information. Before conducting the study, there is rarely enough information about expected effect sizes to adequately assess the required sample size. In this talk, I want to outline a method for continuous research design evaluation that allows for flexible re-adjustments of the sampling plan based on the best available knowledge about parameters at the time. The method is based on projecting Bayes factors into a finite future where sampling variability can still affect results. I will provide several examples for how continuous research design evaluation can increase the informativeness and efficiency of study designs in practice.

Student Discussants: Maximilian Linde and Dandan Tang

Stoevenbelt, Andrea H. (IOPS)

“Time limits as potential source of gender bias in experimental settings: The case of stereotype threat”

Stereotype threat (ST) is often proffered as an explanation of the observed gap in academic performance between Black and White students (Steele & Aronson, 1995), and the gender gap in quantitative domains (Spencer, Steele, & Quinn, 1999). Even though the ST is attested by several meta-analyses, the literature is plagued by publication bias and a failure to consistently replicate results. To address these concerns and to provide an unbiased effect size estimate, I conducted a multilab preregistered replication of the stereotype threat effect on the performance of women in mathematics. Moreover, in the light of the null findings of several preregistered studies, I sought to find an alternative explanation for these inconclusive findings. Flore (2018) concluded that larger stereotype threat effects have been associated with larger proportions of missing data, often caused by the time limits included in stereotype threat experiments. However, this idea has only been studied using total scores, rather than being directly modeled. I study this potential relationship between missing data and the ST effect by applying a two-dimensional item response model that jointly models the relationship between the latent traits mathematics ability and propensity to miss items (Glas & Pimentel, 2008; Glas, Pimentel & Lammers, 2015). I present a novel modeling approach to study stereotype threat, and to assess (gender) bias in experimental settings where tasks are administered under time limits.

Student Discussants: Danielle McCool and Paul Lodder

Stump, Annika (SMiP)

“Is it all about the feeling? Affective and (meta-)cognitive mechanisms underlying the truth effect”

People are more likely to judge repeatedly perceived statements as true. A decisive explanation for this so-called truth effect is that the repeated information can be processed more fluently than new information and that this fluency experience renders the information more familiar and trustworthy. Little is known, however, regarding whether and how affective states and dispositional cognitive preferences influence the truth effect. To this end, we conducted two experiments in which we manipulated (a) processing fluency via repetition, (b) the time interval (10 min vs. 1 week) between repetitions, and (c) short-term affective states using the presentation of emotional faces (Experiment 1) or the presence of an irrelevant source for changes in affective states (Experiment 2). Additionally, we assessed the dispositional variables need for cognitive closure (NCC), preference for deliberation (PD) and preference for intuition (PI). Results of Experiment 1 showed that the truth effect was significantly reduced for statements that were followed by a negative prime, although this was the case only for the longer repetition lag. Furthermore, higher NCC and lower PD scores were associated with an increased truth effect. Results of Experiment 2 replicated the moderating role of NCC and further showed that participants, who were provided with an alternative source for changes in their affective states, showed a reduced truth effect. Together, the findings suggest that (a) fluency-related changes in affective states may be (co-)responsible for the truth effect, (b) the truth effect is decreased when the repetition interval is long rather than short, and (c) the truth effect is increased for individuals with a higher need for cognitive closure. Theoretical implications of these findings are discussed.

Student Discussants: Sebastian Castro Alvarez and Susanne Frick

van den Akker, Olmo R. (IOPS)

“Selective Hypothesis Reporting in Psychology”

Good scientists should be open-minded in the sense that they consider all new evidence, hypotheses, theories, and innovations, even those that challenge or contradict their own interests. However, scientists do not always abide by this Mertonian norm. Studies have shown that researchers often add, drop, or alter study elements when preparing reports for publication, a practice called selective reporting. In this study, we investigated different forms of selective hypothesis reporting (adding hypotheses, omitting hypotheses, promoting hypotheses, demoting hypotheses, and changing hypotheses) by identifying the hypotheses in a large number of preregistration-study pairs. We found that selective hypothesis reporting is widespread, but arguably more important is that we found that hypotheses are often very difficult to identify in preregistrations and papers. In this talk, I will present a selection of vague, incomprehensible, and ambiguous hypotheses that we encountered during data collection. This can hopefully stimulate a discussion about how we can improve hypothesis formulation in psychology.

Student Discussants: Paul Lodder and Richard Artner