



Society for Computers in Psychology  
Annual Convention

*Chicago Hilton*

*November 19, 2015*



## *Message from the President*

Welcome to SCiP 2015. I am excited about this year's SCiP conference, and I hope you are, too. We have a full program including 30 submitted talks, two submitted symposia, and fourteen posters. As well, we have included two events on the theme of "Computational approaches to narrative." The first is a keynote presentation by Patrick Winston, "Understanding story understanding," and the second is the Presidential Symposium, with talks by Mark Finlayson, Menahem Yearia, and Hans-Joachim Backe. As a whole, I think the conference promises to be a stimulating survey of computational approaches to tools, applications, and insights in psychology.



I think that the theme of "Computational approaches to narrative" is appropriate and timely for SCiP. Narrative comprehension was one of the first big challenges of research in artificial intelligence, and the early theorizing of Schank and Abelson and others had a large impact on cognitive psychology and discourse processing in particular. Narrative has continued to be an important problem in artificial intelligence. However, with a few notable exceptions, computational approaches have not figured heavily in more recent research in discourse processing. My thinking is that the time may be right for bringing some of these threads together, and that SCiP might be an appropriate forum to do so. My hope is that the keynote and symposium talks will provide a insightful survey of current ideas and approaches that will enrich the psychological study of narrative.

We also have a technical innovation for this year's conference: a mobile app. To get the conference schedule and information, you need to download the app "Guidebook," and then within the app, search for "SCiP2015" or scan the QR code on the next page. The app should give you access to the conference schedule, lists of speakers and posters, and abstracts. While a conference of this size clearly does not warrant writing a mobile app from scratch, Guidebook (and several other similar apps) allow us to get much of functionality of a dedicated conference app simply by uploading a few files. The Society's steering committee would be interested in your feedback about this approach and how useful it is.

Finally, I would like to remind our conference attendees that the SCiP conference could not occur without the efforts and support from a variety of sources. In particular, I would like to thank our society Secretary/Treasurer, Rick Dale, for his tireless work and organizational acumen in making sure that the many details of running the society and the conference are handled appropriately. I would also like to thank the other members of the conference organizational committee, Pablo Gomez and Shane Mueller, who have provided critical insight and advice on putting the conference together. And of course, I would like to thank the Psychonomic Society, whose financial and organizational support have allowed SCiP conference to flourish as a satellite conference each year.

Thanks for coming to SCiP 2015, and I hope you enjoy the conference!

## *SCiP 2015 Program Committee*

Peter Dixon, University of Alberta (peter.dixon@ualberta.ca)

Rick Dale, University of California, Merced (rdale@ucmerced.edu)

Pablo Gomez, DePaul University (pgomez1@depaul.edu)

Shane Mueller, Michigan Technological University (shanem@mtu.edu)

## *SCiP Officers*

President: Peter Dixon, University of Alberta (peter.dixon@ualberta.ca)

President-Elect: Randall Jamieson, University of Manitoba (randy.jamieson@umanitoba.ca)

Past President: Kay Livesay, Linfield College (klivesa@linfield.edu)

Secretary-Treasurer: Rick Dale, University of California, Merced (rdale@ucmerced.edu)

Federation Representative: Joseph Young (mightyjoephd@verizon.net)

## *Steering Committee*

Zhiqiang (Carl) Cai, University of Memphis (2012-2015)

Monica Riordan, Chatham University (2012-2015)

Todd Gureckis, New York University (2013-2016)

Chris Westbury, University of Alberta (2013-2016)

Scott Crossley, Georgia State University (2013-2016)

Jon Willits, Indiana University (2013-2016)

Stephanie Huetten, The University of Memphis (2014-2017)

Matthew Crump, Brooklyn College of CUNY (2014-2017)

Jennifer Trueblood, Vanderbilt University (2014-2017)

*Get the Guidebook app*

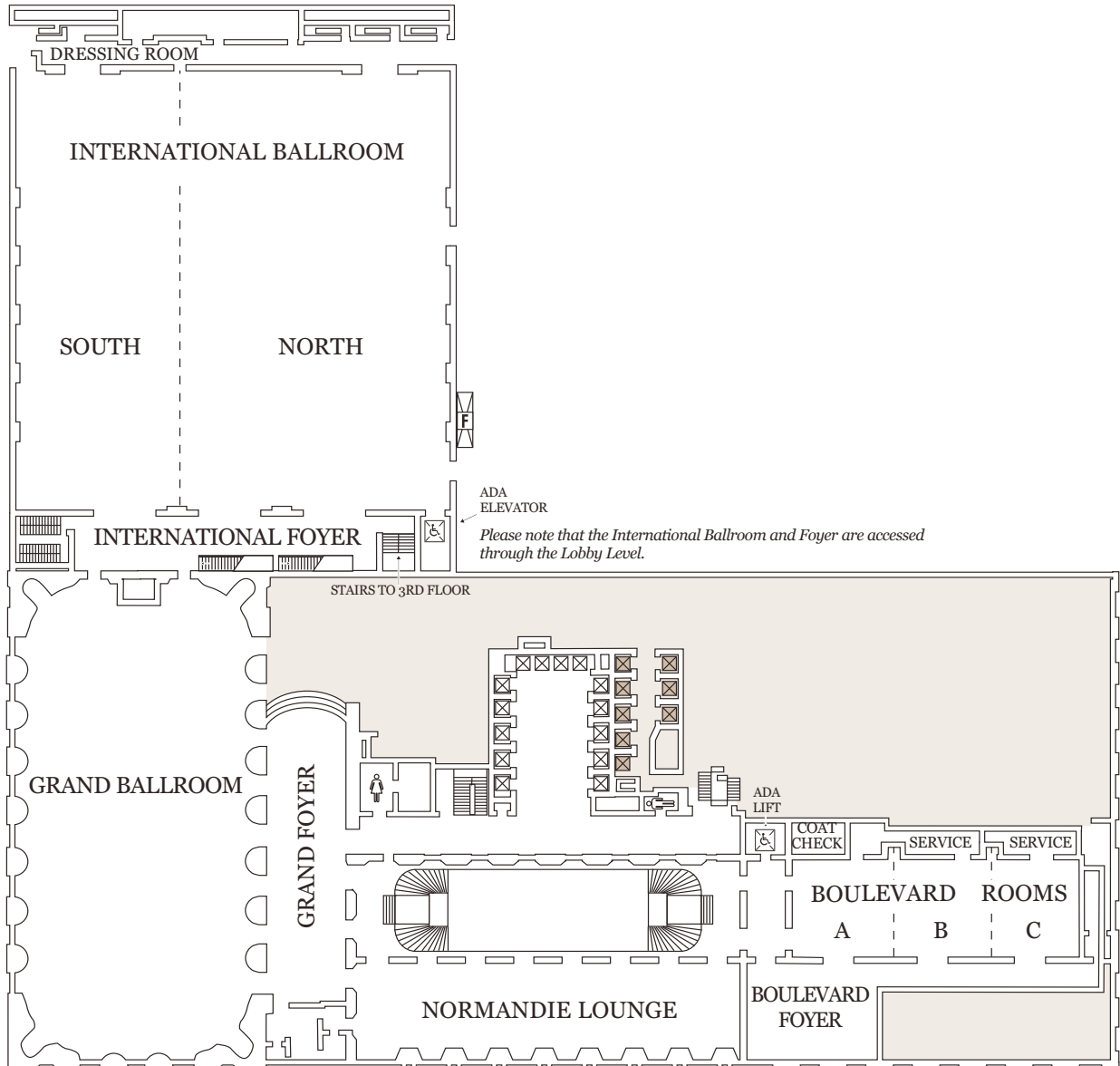


*Get SCiP 2015*

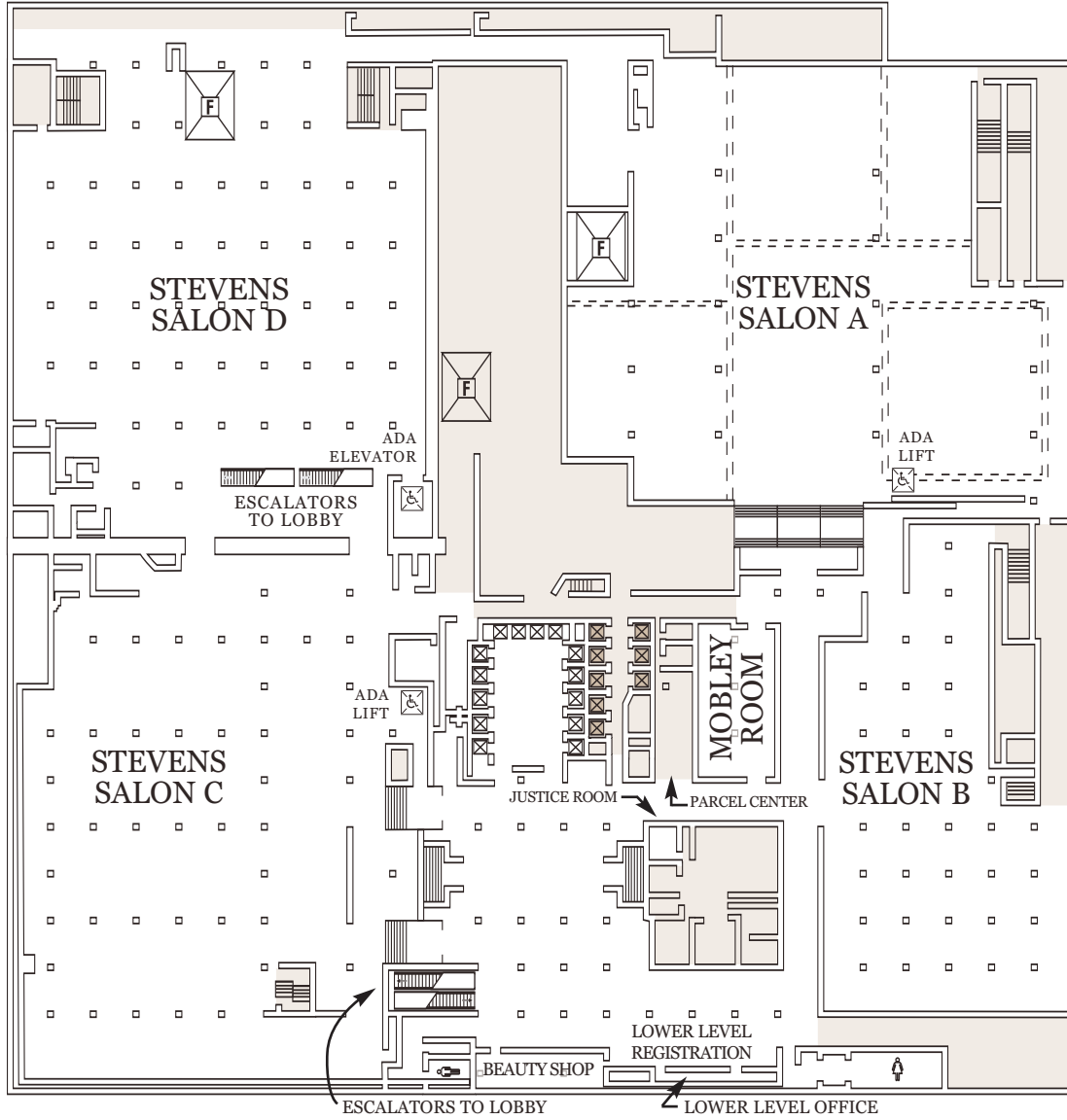


# Floor Plans

## Third Floor (Registration, Breaks, & Talk Sessions)



*Lower Level (Posters)*



## *Information for Presenters*

Digital projectors are available in each of the talk session rooms. Speakers should test their computers and presentation software during the break prior to their talk to ensure that there are no compatibility problems and that they will be able to begin their talk at the scheduled time without delays. Talks should be planned for a duration of twelve minutes, with an additional three minutes allowed for questions. Generally, the last or next-to-last speaker in each session will serve as the moderator and should ensure that each talk begins on time.

Poster presenters should set up their poster in the period from 11:15 to 12:00 and should be available at their poster from 12:15-2:00. Posters should be taken down by 15:00. Posters should be formatted to fit in a 4' × 4' area on half of one side of a poster board. (Note that the area available for a poster is somewhat less than 4' × 4' because of the surrounding frame.) The locations of the posters are numbered as in the program. See <http://www.psychonomic.org/2015-posters> for advice on preparing posters.

## Program Overview

	Boulevard A	Boulevard B	Boulevard C
<b>8:00-8:30</b>	Registration (Boulevard Foyer)		
<b>8:30-9:30</b>	<b>Lexical Representation</b>	<b>Big Data</b>	<b>Modeling &amp; Statistics</b>
	Elizabeth A. Shay	Jason Davis	Drew H. Abney
	Zhiqiang Cai	David W. Vinson	Leslie M. Blaha
	Curt Burgess	Alexandra Paxton	Zach Howard
	Jeff Keith	Timothy Rubin	Henrik Singmann
<b>9:30-9:45</b>	Break (Boulevard Foyer)		
<b>9:45-10:45</b>	<b>Keynote: Patrick Winston (Boulevard C)</b>		
<b>10:45-11:00</b>	Break (Boulevard Foyer)		
<b>11:00-12:15</b>	<b>Language</b>	<b>Measurement</b>	<b>Education</b>
	Nicholas D. Duran	Shane T. Mueller	Wen-Jie Tseng
	Haiying Li	Nada Attar	Stian Reimers
	Colin L. Widmer	Anne Lippert	Dylan Blaum
	Xiaowei Zhao	Sidney D'Mello	E. M. Cedillos-Whynott
	Stephanie Huetten	Patrick Conley	Aiping Xiong
<b>12:15-14:00</b>	<b>Lunch &amp; Posters (Stevens Salon D)</b>		
<b>14:15-15:15</b>	<b>Symposium: Human Ratings</b>	<b>Symposium: Methods Effects</b>	<b>Experimental Tools</b>
	Paweł Mandera	Tim Kuhlmann	Stian Reimers
	Chris Westbury	Michael Dantlgraber	Felix Henninger
	Geoff Hollis	Stefan Stieger	Sébastien Lê
	Jon A. Willits	Ulf-Dietrich Reips	Bonnie Perdue
<b>15:15-15:30</b>	Break (Boulevard Foyer)		
<b>15:30-16:30</b>	<b>Presidential Symposium (Boulevard C)</b>		
	Mark Finlayson		
	Menahem Yearia		
	Hans-Joachim Backe		
<b>16:30-17:30</b>	<b>Business Meeting (Boulevard C)</b>		



## Brief Program

### 8:00-8:30 Registration (Boulevard Foyer)

### 8:30-9:30 Lexical Representation (Boulevard A)

Elizabeth A. Shay , Rajeev D. S. Raizada (University of Rochester), *Using neurobiologically motivated features to investigate the semantic composition of adjectives with nouns*

Zhiqiang Cai, Nia Dowell, Haiying Li, Art Graesser (University of Memphis), *Age of acquisition: How do you remember what words you learned at age one?*

Curt Burgess (University of California, Riverside), *Is the granularity of representations sufficient to account for cerebral asymmetries in semantic priming?*

Jeff Keith, Chris Westbury (University of Alberta), *Heidi-RP: An open-source toolbox for construction, analysis, and exploration of lexical co-occurrence matrices*

### 8:30-9:30 Big Data (Boulevard B)

Jason Davis, David W. Vinson, Rick Dale, Suzanne Sindi (University of California Merced), *The transformative value of high-performance computing to big data research in cognitive science*

David W. Vinson, Jason Davis, Suzanne Sindi, Rick Dale (University of California Merced), *Echoes of social and cognitive constraints on language use in a large naturally occurring data set*

Alexandra Paxton, Rick Dale (University of California, Merced), *Data-driven theory: Using NLP and deep learning in metascientific analyses*

Timothy Rubin, Michael N. Jones (Indiana University), Oluwasanmi Koyejo (Stanford University), Tal Yarkoni (University of Texas Austin), *Neurotopics: Unsupervised learning for discovering function regions of the brain.*

### 8:30-9:30 Modeling & Statistics (Boulevard C)

Drew H. Abney, Ramesh Balasubramaniam (University of California, Merced), *Introduction and applications of the multiscale coefficient of variation (MSCV) analysis*

Leslie M. Blaha (Air Force Research Laboratory), Tim Halverson (Oregon Research in Cognitive Applications, LLC), Brad Reynolds (Wright State University), John T. Balint (George Mason University), *Simcog: Simplified interfacing for modeling cognition*

Zach Howard, Nathan Evans, Scott Brown, Ami Eidels (University of Newcastle), *Extending the detection response task to new domains: A modeling-focused approach*

Henrik Singmann (University of Zurich), Matthew Gretton (University of Tasmania), Scott Brown (University of Newcastle), Andrew Heathcote (University of Tasmania), *rtdists: Distribution functions for accumulator models in R*

**9:30-9:45 Break 1 (Boulevard Foyer)**

**9:45-10:45 Keynote (Boulevard C)**

Patrick Winston (Massachusetts Institute of Technology), *Understanding story understanding*

**10:45-11:00 Break 2 (Boulevard Foyer)**

**11:00-12:15 Language (Boulevard A)**

Nicholas D. Duran (Arizona State University), Riccardo Fusaroli (Aarhus University), Alexandra Paxton (University of California Merced), *Assessing lexical, syntactic, and conceptual turn-by-turn alignment in conversations involving conflict and deception*

Haiying Li , Leah C. Windsor, Arthur C. Graesser (University of Memphis), *Formality and charismatic political leaders: Power and education level*

Colin L. Widmer (Miami University), *A web-based Wizard of Oz system for discourse research*

Xiaowei Zhao (Emmanuel College), *Building bilingual semantic representations based on the Contextual Self-Organizing Map*

Stephanie Huetten, Ariel Mathis (University of Memphis), *Artificial language learning high and low: Representation and response*

**11:00-12:15 Measurement (Boulevard B)**

Shane T. Mueller, Yin-Yin (Sarah) Tan (Michigan Technological University), Ahmed Mohamed (University of Nottingham, Malaysia), Yunier Broche Pérez (University Marta Abreu of Las Villas), *A multi-national norming study for executive function tests*

Nada Attar, Paul Fomenky, Wei Ding, Marc Pomplun (University of Massachusetts, Boston), *An unsupervised time-series learning method for visual search leveraging preprocessed cognitive load pupil data*

Anne Lippert, Justin Wedeking (University of Kentucky), *Measuring expertise in a dynamic domain: Using semantic networks and information density as indicators*

Sidney D'Mello, Robert Bixler (University of Notre Dame), *Automatic gaze-based objective measurement of mind wandering during computerized reading*

Patrick Conley, Jacob Brinkman-Hernandez, Rowan Pockat, Amy Heller, Rachel Ackley (University of Wisconsin, Stevens Point), *Video game genre experiences as predictors of visual ability*

**11:00-12:15 Education (Boulevard C)**

Wen-Jie Tseng, Ching-Fan Sheu (National Cheng Kung University), *Creating Shiny applications for teaching statistical concepts*

Stian Reimers (City University London), *A gamified system for in-lecture SMS text-message quizzes*

Dylan Blaum (Northern Illinois University), Simon Hughes, Peter Hastings (DePaul University), M. Anne Britt, Patty Wallace (Northern Illinois University), Project READi (University of Illinois, Chicago), *Classifying students' causal explanation quality and detecting improvement*

Elizabeth M. Cedillos-Whynott, Christopher R. Wolfe (Miami University), *Assessing the effects of video-games on the development of problem solving and creative thinking*

Aiping Xiong, Weining Yang, Ninghui Li, Robert W. Proctor (Purdue University), *Improving detection of phishing attacks by directing users' attention to domain highlighted URLs*

### **12:15-14:00 Lunch & Posters (Stevens Salon D)**

David A. Washburn (Georgia State University), *A comparison of methods for analyzing individual differences in small-N designs*

Youxuan "Lucy" Jiang, Christopher R. Wolfe (Miami University), *Improving education about women's heart health with a web-based intelligent tutoring system*

Qinyu Cheng, Yehui Liu, Haiying Li, Zhiqiang Cai, Arthur C. Graesser (University of Memphis), *Session recovery of AutoTutor system with rich media client*

Dawn Carroll (California State University, Fullerton), *Technology dependency: Social media and language behavior*

Jon D. Elhai (University of Toledo), Brian J. Hall (University of Macau), *Anxiety about data breaches: Results from a community sample*

Jon D. Elhai, Jason C. Levine (University of Toledo), Brian J. Hall (University of Macau), *Anxiety about electronic data breaches and relations with digital privacy protection*

Adrian Cunningham, Curt Burgess (University of California, Riverside), *Life beyond semantic distance using HAL and HiDEx*

Curt Burgess (University of California, Riverside), Justin Estep (San Diego Mesa College), Sarah Maples (University of California, Riverside), *Semantic and grammatical characteristics of closely related verb types*

Benjamin Schloss, Ping Li (The Pennsylvania State University), *Predicting human brain activation for concrete nouns: Comparing vector space and graph theory models*

Melissa Ray, Joe Magliano (Northern Illinois University), *The automatic assessment of moment-to-moment comprehension processes*

David Allbritton (DePaul University), *Making text analysis accessible with R and Shiny*

Evgeny Novikov, Irina Vakoliuk, Igor Varchak, Roman Akhapiin, Dariya Shvaiko, Ekaterina Budenkova, Irina Shalaginova (Baltic Federal University Immanuel Kant), *Computer oculoigraphy system for central nervous system monitoring using standard digital camcorders*

Toshio Ohyanagi, Kunihiro Kanaya, Yasuhito Sengoku (Sapporo Medical University), Ada Leung, Lili Liu, Masako Miyazaki (University of Alberta), *Development of a new Bluetooth Low Energy device for measuring accurate reaction time*

Josh de Leeuw (Indiana University), *A collaborative, open-source collection of browser-based experiments for teaching demonstrations using jsPsych*

Janelle Szary, Michael N. Jones (Indiana University), Rick Dale (University of California, Merced), *Model selection using multi-model inference: An example showing interaction dominance in collaborative foraging*

Melody W. Dye (Indiana University), Daniel Yarlett (Stanford University), Michael Ramscar (Tübingen University), Michael N. Jones (Indiana University), *The temporal dynamics of distributional semantics*

**14:15-15:15 Symposium: Extrapolating human ratings using distributional semantics and machine learning methods: Advances and limitations (Boulevard A)**

Paweł Mandra, Emmanuel Keuleers (Ghent University), *Predicting item-level effects of relatedness with models based on prediction and counting*

Chris Westbury, Elena Nicoladis, Jeff Keith (University of Alberta), *Predicting semantics on the fly: Affect-based computational estimates of noun gender judgments in unilingual English speakers*

Geoff Hollis, Chris Westbury (University of Alberta), *A computational modelling approach to understanding human color preferences*

Jon A. Willits (University of California Riverside), Tim Rubin, Michael N. Jones (Indiana University, Bloomington), *If you want to know which semantic model is best, don't compare semantic models*

**14:15-15:15 Symposium: Methods effects revisited: Current state of the art of personalization, Radex models, seriousness checks, and visual analogue versus Likert scale format in web-based research (Boulevard B)**

Tim Kuhlmann, Michael Dantlgraber, Ulf-Dietrich Reips (University of Konstanz), *Advantages of visual analogue scales for brief internet-based measurements*

Michael Dantlgraber, Ulf-Dietrich Reips (University of Konstanz), *A new way of factor structure visualization: The Factor Based Radex Model*

Stefan Stieger, Ulf-Dietrich Reips (University of Konstanz), *Impact of personalization of invitation letters to online studies: Study duration and outcome feedback*

Ulf-Dietrich Reips (University of Konstanz), *Revisiting the seriousness check: A simple method to increase data quality in internet-based research*

**14:15-15:15 Experimental Tools (Boulevard C)**

Stian Reimers (City University London), Neil Stewart (University of Warwick), *Audio presentation accuracy in Adobe Flash and HTML5/JavaScript Web experiments*

Felix Henninger (University of Koblenz-Landau, Max Planck Institute for Research on Collective Goods, University of Mannheim), Pascal J. Kieslich (University of Mannheim), *Psynteract: A flexible, cross-plattform, open solution for interactive experiments*

Sébastien Lê, Tâm Minh Lê (Agrocampus Ouest), *Holos: A collaborative environment for holistic approaches*

Bonnie Perdue (Agnes Scott College), *Computers and comparative cognition research at the zoo*

**15:15-15:30 Break 3 (Boulevard Foyer)**

**15:30-16:30 Presidential symposium: Computational approaches to narrative (Boulevard C)**

Mark Finlayson (Florida International University), *Computing Bartlett: A scientific approach to narrative analysis*

Menahem Yearia (Bar Ilan University), Paul van den Broek (Leiden University), *A computational modeling of semantic knowledge in reading comprehension: Integrating the Landscape Model with the latent semantic analysis*

Hans-Joachim Backe (IT University of Copenhagen), *The glove and the hand: A Lacanian approach to player-avatar relationship*

**16:30-17:30 Business Meeting (Boulevard C)**

## Detailed Program

### 8:00-8:30 Registration (Boulevard Foyer)

### 8:30-9:30 Lexical Representation (Boulevard A)

- 8:30 Elizabeth A. Shay, Rajeev D. S. Raizada (University of Rochester), *Using neurobiologically motivated features to investigate the semantic composition of adjectives with nouns*

Language is not just isolated words; there are interactions between the meanings. In the current work, we used neurobiologically-based features of meaning (sensory-motor features: sound, color, manipulation, motion and shape) to compare two popular models of this process of semantic composition: element-wise vector addition and multiplication. A large literature (e.g. Binder's work such as Fernandino et al., 2015) suggests that perceptual systems contain information that can be extracted using neural decoding (e.g. Anderson, Murphy & Poesio, 2014). Within semantic composition research, adjective-noun composition has been relatively well-studied due to their concrete nature, making this ideal for exploration of composition of these neurobiologically-based features. Using Amazon's Mechanical Turk, participants rated how much each of the words and phrases (made of all combinations of the selected adjectives and nouns) evoked the features. Although there are several potential composition functions suggested in the literature, the two most successful, easily interpretable functions have been multiplication (e.g. Chang et al., 2009) and addition (e.g. Mitchell & Lapata, 2010). Both multiplication and addition surpass chance at matching the correct phrase, but addition outperformed multiplication (addition = 7.6/60, multiplication = 13.4/60). Addition allows the adjective to weight the important sensory-motor attributes for the noun. Based on these behavioral results, we predict, and will test in upcoming work, that addition will also be successful when using brain activity (from fMRI) as the representations of the adjectives, nouns and phrases.

- 8:45 Zhiqiang Cai, Nia Dowell, Haiying Li, Art Graesser (University of Memphis), *Age of acquisition: How do you remember what words you learned at age one?*

Age of acquisition (AoA) has been identified as a lexical feature that affects human's processing speed and accuracy in reading. Researchers have investigated how AoA influences psychological processes. During the last decade, AoA has been integrated into "Coh-Metrix," a widely used text analysis tool for researchers. Research on AoA has typically been based on the human rated norms. The ratings of AoA have been collected by presenting lexical items to human subjects and asking the subjects to report the age at which the items were learned. The latest technology used in collecting the ratings is the web-based crowdsourcing from Amazon Mechanical Turk. There are interesting questions to ask about such ratings. How can a human subject of a certain age, say, 25, remember that an item was learned at an early age, such as 1, 2, or 3? What makes a subject believe that an item was learned at the age he/she reported? To answer such questions, we used lexical features extracted from a large corpus to simulate human's judgments. The results show that the distribution of a word across texts of different grade levels in a large corpus predicts humans' judgments very well. This suggests that (1) texts of different grade levels naturally reflect AoA features and/or (2) humans' judgments about AoA were based on the probability of a word appearing in different grade levels.

- 9:00 Curt Burgess (University of California, Riverside), *Is the granularity of representations sufficient to account for cerebral asymmetries in semantic priming?*

There has been a long-standing debate, especially in neurolinguistics, about how best to characterize the nature of semantic representations and the type of processing that accounts for cerebral asymmetries in semantic priming. On one hand cerebral asymmetries in the priming of ambiguous words can be described as a function of the amount of time the subject has processed the prime. The time course of activation for dominant meaning is rapid in the left-hemisphere (LH) and is maintained (depending on sentential needs) whereas the subordinate activation is slower and inhibited in the LH. In the right-hemisphere, both meanings stay activated until higher-level processing makes a commitment (Burgess & Simpson, 1988, Burgess & Lund, 1998). The time-course model has been contrasted with an alternative account that is representational in nature. Beeman (1998; Jung-Beeman, 2005) has argued that the hemispheric asymmetries can be traced to the LH semantic system using a finely-coded representation and the RH system using a coarsely-coded representation. Coarsely-coded meanings have fewer associates or features that convey less meaning information. In this study, the HAL model was used to operationalize the notion of fine and coarse coding by manipulating the amount of information in the meaning vectors. Coarsely-coded representations would have fewer vector elements where the dimensional reduction forces more meaning information into fewer representational elements. The type of semantic distance (city-block or Euclidean) was also manipulated. Results show that the number of vector elements had a negligible effect on priming.

- 9:15 Jeff Keith, Chris Westbury (University of Alberta), *Heidi-RP: An open-source toolbox for construction, analysis, and exploration of lexical co-occurrence matrices*

Word co-occurrence models of semantics bootstrap lexical semantics from patterns of word co-occurrence. Several implementations of word co-occurrence models exist, including at least one open-source model, Shaoul & Westbury's HiDEx, after which our new model Heidi is named. Heidi is implemented in R, with associated tools written in Python (hence, Heidi-RP). Implementing a co-occurrence model in R and Python makes it easier to bring a wide set of open source functions in both languages to the analysis and exploration of the structure of co-occurrence space. In this presentation we will discuss our analysis of the relationship between novel semantic neighborhood density measures (e.g., entropy measures of a target word's distribution of cosine distances between itself and all other words in the lexicon; neighborhood counts based on various levels of standard deviation cutoffs taken from a target word's standardized distribution of cosine distances between itself and all other words in the lexicon; etc.) and human lexical access and semantic decision measures.

### 8:30-9:30 Big Data (Boulevard B)

- 8:30 Jason Davis, David W. Vinson, Rick Dale, Suzanne Sindi (University of California Merced), *The transformative value of high-performance computing to big data research in cognitive science*

The growth of large, naturally occurring datasets is eclipsing the ability of existing tools to efficiently uncover behavioral phenomena. These tools have not kept pace with new computing technologies, which require explicit consideration of many cores and heterogeneous compute devices. Languages like Python and R, while renown for their ease of use and powerful, high-level language features, are not yet able to take full advantage of this new hardware. This has

created an opportunity for computer scientists to edge in and write hardware-optimized code that implements high-performance, statistically-driven models that often lack the sophisticated behavioral theories that have emerged over the last 50 years of experimental psychology. To restore the computational social scientist's ability to apply these theories to emerging datasets, we are developing a library of algorithms that solve a standard suite of computational problems frequently encountered in behavioral analysis. We aim to do so in a scalable way by exploiting modern high-performance computing (HPC) techniques and hardware. This library will then provide high-level bindings to Python and R. As a proof of concept, we focus on the use of basic n-gram models to analyze a Yelp, Inc. dataset, which was previously intractable on commodity hardware with current NLP libraries in R. We provide benchmarks contrasting canonical approaches with a novel data structure we developed and demonstrate that it enables an entirely new scale of n-gram analyses. This demonstration will provide attendees with an introduction to the analysis workflow in HPC, and a concrete example in the analysis of large amounts of linguistic data.

8:45 David W. Vinson, Jason Davis, Suzanne Sindi, Rick Dale (University of California Merced), *Echoes of social and cognitive constraints on language use in a large naturally occurring data set*

Some recent theories of language see it as a complex and highly adaptive system, adjusting to factors at various time scales. For example, at a longer time scale, language may adapt to certain social or demographic variables of a linguistic community. At a shorter time scale, patterns of language use may adjust to the cognitive affective states of the language user. If so, language may be used as an indication of certain cognitive and social influences on behavior. Specifically, it may reveal subtle "echoes" of constraints that reflect these adaptive processes. These echoes may only be detectable with adequate data to identify their presence. Until recently, data sets large enough (and computational tools efficient enough) to test how subtle effects of socio-cultural properties—spanning vast amounts of time and space—influence language change have been difficult to obtain. We analyzed over one million online business reviews using information theory and network analyses to quantify language structure and social connectivity while review ratings provide an indication of the reviewer's cognitive affective state. Results indicate that some proportion of variance in the structure of individual language use can be accounted for by differences in cognitive states and social-network structures, even after fairly aggressive covariates have been added to regression models.

9:00 Alexandra Paxton, Rick Dale (University of California, Merced), *Data-driven theory: Using NLP and deep learning in metascientific analyses*

Recent advances in natural language processing (NLP) and deep learning tools – coupled with increasingly large amounts of freely available linguistic data – can provide insights about language use on a previously unimagined scale. In addition to analyzing patterns of everyday language use in digital communication, these methods can also be used to answer metascientific questions by analyzing scientific communication. While most previous metascientific applications to date tend to describe scientific activity broadly (e.g., Griffiths & Steyvers, 2004), we believe these tools may also provide a data-driven approach to analyzing the relations among scientific theories. As an example, we present an application of these methods to understand the relations among various theoretical terms in the research area of interpersonal coordination (also known as synchrony, mimicry, alignment, contagion, and more). We use automated language analysis tools on a corpus of thousands of scientific abstracts to quantify the relationships among these terms and their underlying theoretical perspectives, which are often blurred in discussions within this research area. Such analyses uncover implicit associations within the literature and shed light on researchers' perspectives on



these terms. This work demonstrates how researchers may use NLP and deep learning to refine theoretical perspectives and associations through metascientific analysis.

- 9:15 Timothy Rubin, Michael N. Jones (Indiana University), Oluwasanmi Koyejo (Stanford University), Tal Yarkoni (University of Texas Austin), *Neurotopics: Unsupervised learning for discovering function regions of the brain.*

We used a class of novel unsupervised Bayesian models based on Latent Dirichlet Allocation (Blei, Ng & Jordan, 2003) to identify and discover functional brain regions in the Neurosynth meta-analytic database of over 9,700 functional MRI studies. Our model takes as input a corpus of unlabeled fMRI publications in which the abstract text and the reported (x,y,z) activation coordinates for each publication have been extracted. From this input, the model learns a set of T "Topics", where each topic is parameterized by a probabilistic distribution over linguistic features (e.g. vision, motion, perception) and a spatial distribution (in this case, a 3-dimensional Gaussian distribution). Intuitively, each of these topics captures a functional region of the brain, where the linguistic features for the topic describe a functional signature associated with the spatial distribution of the topic. This approach has significant advantages over alternative methods – such as those based on Principal Components Analysis (PCA) – in that it does not treat voxels as spatially independent. This encourages more interpretable parcellations that respect known local correlations between brain physiology and cognitive function. When trained on the entire Neurosynth database, our model produces highly interpretable results that both replicates known associations and suggests novel hypotheses suitable for further investigation.

### 8:30-9:30 Modeling & Statistics (Boulevard C)

- 8:30 Drew H. Abney, Ramesh Balasubramaniam (University of California, Merced), *Introduction and applications of the multiscale coefficient of variation (MSCV) analysis*

Linear and nonlinear time series analyses afford quantification of distributional, sequential, and multiscale properties for behavior that varies over time. However, not all existing time series analyses allow reliable parameter estimation for time series that are of extremely short lengths, e.g.,  $\leq 10$  data points. Current time series analyses amenable to short time series only capture linear, sequential information, but much work in the past few decades has suggested that the multiscale structure of cognitive and behavioral phenomena can provide important insights, too. The current talk will introduce a novel analysis that can estimate the multiscale structure of time series that greatly vary in length. The multiscale coefficient of variation (MSCV) analysis measures the distance between local coefficient of variation estimates at particular timescales and the overall coefficient of variation across the entire time series. The user identifies the range of timescales of interest. For each timescale, coefficient of variation is estimated across a sliding window from the start to the end of the time series. Average coefficient of variation values at each timescale are summed, normalized by the overall coefficient of variation, and then divided by the total number of timescales. We will provide a short introduction of how the MSCV works using simulations from ARFIMA models. We will then show applications of the MSCV analysis to a wide array of datasets, spanning from musical corpora to a multi-language database of speech rhythms to center-of-pressure estimates of postural sway.

8:45 Leslie M. Blaha (Air Force Research Laboratory), Tim Halverson (Oregon Research in Cognitive Applications, LLC), Brad Reynolds (Wright State University), John T. Balint (George Mason University), *Simcog: Simplified interfacing for modeling cognition*

Two continuing hurdles for cognitive modeling in psychology are the difficulty with enabling models to interact with the same tasks as the user and the challenge of understanding the internal dynamics of computational models during simulation. Multiple attempts have been made to add this functionality to computational cognitive architectures, but no comprehensive solution yet exists. Our research presents a solution that enables computational models to interact with web browser-based software while requiring little modification to the task code. Simplified Interfacing for Modeling Cognition - JavaScript (SIMCog-JS) allows the modeler to specify how elements in the interface are translated into model-parsable information and allows events to be sent from the model to the task and other external software. SIMCog-JS also facilitates model understanding by displaying the internal workings of a computational cognitive model in a visualization dashboard and embodying model behavior in an animated agent. For the latter, SIMCog-JS passes the model events and internal state information which are translated into responses in a virtual task interface. Model events and states are stored to a file that is used in a browser-based visualization dashboard to illustrate the internal model activity for debugging and evaluation. We demonstrate SIMCog-JS with the computational cognitive architecture ACT-R performing dynamic visual multitasking in a JavaScript environment containing 1, 2 or 4 simultaneous tasks. The model visualizations demonstrate the internal competition for perceptual and motoric resources. Future capabilities look to interface SIMCog with additional modeling formalisms and to incorporate real-time performance analytics.

9:00 Zach Howard, Nathan Evans, Scott Brown, Ami Eidels (University of Newcastle), *Extending the detection response task to new domains: A modeling-focused approach*

One of the major "real-world" contributions of cognitive psychology in recent times has been the finding that using a mobile phone impairs driving performance. Studies applying typical stimulus-response reaction time measures (referred to as the Detection Response Task; DRT) have demonstrated that mobile phone conversations increase "cognitive load" to such an extent that driving performance suffers noticeably. These findings have led to the introduction of legislation in many countries prohibiting the use of mobile phones whilst driving. Although the focus of this research has been on driving scenarios, the DRT measure has scope to be applied in a much broader domain of research. Indeed, the relative simplicity and convenient application of the measure makes it somewhat uniquely suited to be used in applied, or industry settings, as we can collect informative data about "cognitive load" outside of the typical lab setting. In the current presentation, we detail our lab's initial and planned future applications of the DRT with industry partners in areas as diverse as the aeronautical and coal-mining sectors, and demonstrate how the psychological concept of "cognitive load" has significant real-world meaning in these high-stress industries. We also outline how the same measure can be used to both inform and enhance the practices of such diverse industries. Finally, we discuss the potential application of cognitive modeling techniques, such as Systems Factorial Technology, or accumulator fitting procedures to DRT data to obtain novel estimates of cognitive mechanisms in real-world tasks.

9:15 Henrik Singmann (University of Zurich), Matthew Gretton (University of Tasmania), Scott Brown (University of Newcastle), Andrew Heathcote (University of Tasmania), *rtdists: Distribution functions for accumulator models in R*

Accumulator models for jointly modeling both decisions and their response times provide one of the most widely used modeling framework in cognitive psychology. The primary exemplars are the standard diffusion model with trial to trial variability in bias, evidence accumulation rates and non-decision time, which applies to two-choice data, and the standard linear ballistic accumulator (LBA), with the same types of trial to trial variability, which applies to choices among any number of alternatives. We present an R package, *rtdists*, providing for the first time the basic functions required to simulate and estimate these models through maximum likelihood, Bayesian, and quantile-based methods; probability density functions, cumulative density functions, and random number generation functions. For the LBA we also provide this functionality for new versions with four positive drift rates distributions: truncated normal, log-normal, Gamma, and Frechet. For all models the functions are fully vectorized allowing a different parameter value for each data point/accumulator, with the exception of the non-decision time variability in the LBA. We provide examples of how to use *rtdists* in applications.

**9:30-9:45 Break 1 (Boulevard Foyer)**

**9:45-10:45 Keynote (Boulevard C)**

Patrick Winston (Massachusetts Institute of Technology), *Understanding story understanding*

Much of Psychology has been about characterizing human behavior. Much of Artificial Intelligence has been about implementing systems that shed no light on human behavior. I believe the time has come to bring the two together to develop a computational account of human intelligence. I am optimistic, in part, because we know more. Perhaps more importantly, some of us in AI are asking better questions, informed by research ranging from Paleanthropology to Brain and Cognitive Science. In particular, my Genesis Group asks what it is that makes human intelligence different from that of Orangutans or Neanderthals. Our answer is that we are the great story tellers, and believing that, being engineers, we are driven to exploratory programming aimed at better understanding and modeling aspects of our human story competence. Given computer parsable, 100-sentence stories drawn from Shakespearean literature to Crow creation myths, our Genesis System answer simple questions, reads with cultural bias, retells persuasively, teaches with student awareness, exhibits a kind of self-awareness, and raises new questions.

**10:45-11:00 Break 2 (Boulevard Foyer)**

**11:00-12:15 Language (Boulevard A)**

- 11:00 Nicholas D. Duran (Arizona State University), Riccardo Fusaroli (Aarhus University), Alexandra Paxton (University of California Merced), *Assessing lexical, syntactic, and conceptual turn-by-turn alignment in conversations involving conflict and deception*

Our goal is to present a novel systematic approach for automatically assessing conversational alignment across turn-by-turn exchanges, examining linguistic behavior at increasing levels of abstractness through lexical, syntactic, and conceptual similarity. Using the latest advances in Python-based NLP tools, the procedure begins by taking conversational partners' turns and converting each into a lemmatized sequence of words, assigning part-of-speech tags and computing high-dimensional semantic vectors per each utterance. Words and part-of-speech tags are further sequenced into n-grams of increasing length (from uni- to quad-grams) to allow a range of linguistic structures to be examined. Lexical, syntactic, and conceptual alignment values are then calculated on a turn-by-turn basis as cosine scores. To showcase our approach, and to demonstrate its effectiveness in capturing turn-level linguistic alignment, we turn to a unique conversational context: one in which participants disagree or agree with each other about contentious sociopolitical topics, with the added element of one partner secretly taking a "devil's advocate" position. Our findings reveal that high-level intentional factors can modulate alignment processes consistently across multiple levels of linguistic abstraction. Contrary to previous findings on non-verbal coordination (Duran and Fusaroli, under review), deception disrupts verbal alignment, and alignment generally decreases over time. Moreover, for a subset of lexical and syntactic measures, this decrease is most pronounced in truth conversations. This suggests that a hypothesized role of alignment, whereby mutual understanding is facilitated, is established early and is required less as truth conversations progress. Implications for current models of interpersonal coordination will be discussed.

- 11:15 Haiying Li , Leah C. Windsor, Arthur C. Graesser (University of Memphis), *Formality and charismatic political leaders: Power and education level*

This paper conducted an automated text analysis of formality of four charismatic leaders each of whom is the core generation of the Chinese collective leadership of the Communist Party of China (CPC) from the year of 1921 to 2012 across almost a century (N = 1662): Mao Zedong (N = 1158), Deng Xiaoping (N = 221), Jiang Zemin (N = 203), and Hu Jintao (N = 80). We investigated whether the leaders changed their language formality with the corresponding social contexts, such as power and the nation's education level. Formality is a multilevel textual indicator of language style, consisting of the features of language, discourse, and psychology. In this study, formality was measured by linguistic and psychological features at the multitextual levels. A mixed-effects modeling with R (lme4) was used to build two models (power and education). In both models, formality was used as the dependent variable and leaders as the random effect. In the power model, power, age, and power  $\times$  age were used as the independent fixed effect variables. In the education model, education, age, and education  $\times$  age were used as the independent fixed effect variables. Results revealed these leaders' formality increased when they were in power and when the nation's education level was high. In addition, leaders tend to use more informal language when they grow old. The interaction with age reduced leaders' formality. The study implies leaders' formality is related to power and the nation's education and varies over time related to social change.

11:30 Colin L. Widmer (Miami University), *A web-based Wizard of Oz system for discourse research*

Wizard of Oz experiments offer researchers interested in human-computer interactions and discourse a research strategy for assessing dialogue between a user and a computer system that is both practical and capable of answering theoretically motivated questions. Using a human researcher to select dialogue moves "behind the scenes" enables researchers without access to intelligent systems capable of natural language processing to engage in human-computer interaction research and investigate questions about discourse more broadly. To conduct Wizard of Oz experiments on discourse I developed a web-based system that allows users to engage in dialogue with avatars in what appears to be an intelligent computer system. Users enter text responses and then receive responses from the animated talking avatar that are actually dialogue moves selected by a human researcher from a predefined list. This system can be used to quickly and easily design and implement studies that address many theoretically motivated propositions about discourse, from evaluating the impact of specific changes to dialogue (such as formality, cohesion, or affect) to simulating the context of different types of discourse to evaluate user response to dialogue (such as evaluating the impact of different kinds of feedback on student learning or which responses best help people to form a well-developed argument). I will discuss the development and use of the system in research and present preliminary data from a study that uses the system to assess the impact of feedback on user learning in the context of an intelligent tutoring system exemplifying features of the system.

11:45 Xiaowei Zhao (Emmanuel College), *Building bilingual semantic representations based on the Contextual Self-Organizing Map*

Zhao, Li, and Kohonen (2011) recently introduced a software package (Contextual Self-Organizing Map) that applies a corpus-based statistical learning algorithm to derive semantic representations of words. The algorithm relies on the analyses of contextual information extracted from a text corpus, specifically, analyses of word co-occurrences in a large-scale electronic database of text. It also has the capacity to integrate realistic perceptual and sensorimotor features into the semantic representations of words. In the current study, we applied the program to a bilingual situation (English and Chinese). Particularly, we examined how the semantic structure of L2 words can be built based on and influenced by the semantic representations of L1 words in a sequential L2 learning situation. We derived the semantic representations under two conditions: with or without integrating the perceptual and sensorimotor features of certain words. The results were processed and illustrated on a self-organizing map, an unsupervised neural network model that projects the statistical structure of the context onto a 2-D space.

12:00 Stephanie Huetten, Ariel Mathis (University of Memphis), *Artificial language learning high and low: Representation and response*

In real world situations, infant and adult language learners alike will encounter negation such as "No, not the red one" or "Don't touch that". Behavioral consequences of such contexts include looking for alternatives or stopping actions in progress. The learning process of negation in language is not yet well understood. To investigate this, an artificial language learning paradigm was developed with prefixes on nouns denoting "yes, pick this" or "no, pick the other". Participants sat at a computer and heard a word with prefix, then chose between two pictures of familiar objects. Feedback on accuracy was immediately given. Pairs of items were yoked, and either always Negated, always Affirmative, or Both affirmative and negated. A testing phase using three pictures and omitting prefixes assessed learning. Currently, six people have participated. One was discarded due to self-reported brain injury. Performance across all

training showed lowest performance for items trained as Both Affirmative and Negated ( $M=68\%$ ,  $SD=14\%$ ), second lowest for Negated items ( $M=78\%$ ,  $SD=16\%$ ) and highest for Affirmative ( $M=87\%$ ,  $SD=6\%$ ). During the testing phase, one participant scored 100%, and two participants interpreted the Negated trained items as the alternative yoked item. For example “an-sulto” meant “not the bird” and they clicked on the house. They interpreted the word stem “sulto” as meaning house. The response behavior thus dominated over the linguistic meaning for these participants indicating potentially interesting individual differences in response behavior dominance and language representation. This study is currently in progress and statistical inferences will be made with more participants.

### 11:00-12:15 Measurement (Boulevard B)

11:00 Shane T. Mueller, Yin-Yin (Sarah) Tan (Michigan Technological University), Ahmed Mohamed (University of Nottingham, Malaysia), Yunier Broche Pérez (University Marta Abreu of Las Villas), *A multi-national norming study for executive function tests*

Performance on cognitive and computer-based testing often differ across nationality and culture, but relatively few free and open source tests have been systematically assessed on participants in different nations to assess whether normative performance is similar. Substantial differences on some tasks (including executive function) might be expected because of differences in education, bilingualism, and east/west cognitive styles. To help identify systematic similarities and differences across cognitive tests, we developed a battery of open source tests available within the PEBL Test Battery, and collected data at multiple international sites. We report the results of data collected in Malaysia, Taiwan, United States, and Cuba on six cognitive tests (Corsi blocks, number Stroop, Tower of London, Tower of Hanoi, Eriksen Flanker, and color Stroop). Results indicate broad similarities in test performance across cultures, with some systematic differences mainly stemming from Cuban participants who were generally slower and more accurate than participants at other sites. We discuss the general lessons learned review systematic differences between participants.

11:15 Nada Attar, Paul Fomenky, Wei Ding, Marc Pomplun (University of Massachusetts, Boston), *An unsupervised time-series learning method for visual search leveraging preprocessed cognitive load pupil data*

Measurement of a user’s cognitive load (CL) is important for developing and evaluating user interfaces. Any experimental session that collects data from human subjects using a noninvasive monitoring system such as eye-tracking poses a challenge to pattern recognition due to the noise in the data and the large feature-to-instance ratio. This paper investigates the suitability of the Random Subspace (RS) ensemble method using sequential window frames of different sizes over the time of the experiment to predict the level of CL for two different visual search tasks based on only one feature (pupil size). We also use RS to preprocess pupil data - filtering out data not reflecting cognitive workload. The method reported classification accuracies up to 85% for detecting the levels of cognitive load comparing to the conventional approach which was 66%. In addition, our study shows that using the RS method, eye tracking data can effectively be preprocessed in a way that removes samples that do not reflect CL. Using this method, it is found that samples are evenly dropped from both conditions, producing balanced datasets. The RS preprocessing method kept 85% of the original dataset, compared to 68% for conventional preprocessing. We have shown that machine learning techniques can be applied to classify different levels of CL from visual task performance. The study also provides an effective method

to reduce error related to preprocessing of pupil data before analysis, and includes as much data as possible in the analysis

- 11:30 Anne Lippert, Justin Wedeking (University of Kentucky), *Measuring expertise in a dynamic domain: Using semantic networks and information density as indicators*

Traditional measures of expertise assume the domain space of the expert is stable. However, unlike domains such as music, tennis or chess, the legal domain, like many others, is dynamic and becoming increasingly complex. This means traditional measures of expertise will not be sufficient to accurately measure the success with which individuals perform legal tasks. Instead, a new, dynamic measure of expertise is necessary for the legal domain, and we draw upon methods from cognitive psychology and information science to create such a measure. Specifically, we examine U.S. Supreme Court opinions from 20 justices over their careers in three issue areas. We propose a measure of expertise that incorporates the network coherence of a text and also a computational measure of a text's information density (Review-Internal Entropy). We find that some justices were able to show signs of increased expertise while others are not able to do so.

- 11:45 Sidney D'Mello, Robert Bixler (University of Notre Dame), *Automatic gaze-based objective measurement of mind wandering during computerized reading*

Mind wandering is a ubiquitous phenomenon where attention involuntarily shifts from task-related thoughts to internal task-unrelated thoughts. The last decade has witnessed an explosion of research into the antecedents, behavioral correlates, and consequences of mind wandering. Unfortunately, the scientific study of mind wandering has been stymied by a lack of objective measures, leading to a near exclusive reliance on self-reports. We investigated the use of eye gaze to automatically measure mind wandering during a laboratory computerized reading task. Participants (N = 178) were pseudo-randomly probed to report mind wandering while a Tobii TX300 eye tracker recorded their gaze during reading. Supervised machine learning methods (specifically Bayesian classifiers) were used to automatically discriminate between the positive (30%) and negative (70%) responses to the probes in a manner that generalized to new participants. The mind wandering detection accuracy was 72% (expected accuracy by chance was 60%) when probed at the end of a page and 67% (chance was 59%) when probed in the midst of reading a page. Global gaze measures (gaze patterns independent of content – e.g., mean fixation duration) were more effective than content-specific local gaze features (e.g., first pass fixations). The automatically detected mind wandering rate was significantly negatively correlated with measures of learning (Spearman rhos of -.361 and -.299) and transfer (Spearman rhos of -.187 and -.235) after controlling for prior knowledge. Thus, for the first time, a fully automated objective measure might be available for laboratory studies of mind wandering during reading.

- 12:00 Patrick Conley, Jacob Brinkman-Hernandez, Rowan Pockat, Amy Heller, Rachel Ackley (University of Wisconsin, Stevens Point), *Video game genre experiences as predictors of visual ability*

Video games have become ever more ubiquitous in our society, with a presence in not just computers and consoles, but now phones, tablets, and even children's toys. In a previous study, we determined that the amount of video game experience and the age at which that experience was gained were predictors in a visual search task, with more and earlier experience leading to faster performance. In the current study, we have examined the role that video game genre plays in this relationship. A questionnaire was distributed asking participants to rate their

amount of experience in thirteen different genres of video game, such as first person shooters, sports games, strategy games, and role playing games. Results demonstrated that experience playing faster-paced or more action-oriented genres of video games led to faster and more accurate performance in a visual reaction time task. Additionally, experience in a greater number of differing genres (that is, exposure to a higher number of different categories of video games) also led to better performance. These results suggest that the type of video gaming experience matters as much as the amount of the experience. The data on what genres are being most played and the attitudes of players towards video gaming in general will also be presented.

### 11:00-12:15 Education (Boulevard C)

- 11:00 Wen-Jie Tseng, Ching-Fan Sheu (National Cheng Kung University), *Creating Shiny applications for teaching statistical concepts*

This paper demonstrates how to implement interactive graphics in "Shiny," a web application with RStudio, to illustrate statistical and psychological models. Three pedagogical examples are chosen: relating the minimization of the sum of squares to the fitted regression lines, moving from priors to posterior distributions as data accumulate in Bayesian inference, and sweeping out ROC curves as criterion changes for signal detection theory models. In addition to providing example scripts for these web applications, we also discuss advantages of embedding interactive graphics in Shiny for teaching and learning statistical and psychological modeling.

- 11:15 Stian Reimers (City University London), *A gamified system for in-lecture SMS text-message quizzes*

I present a new learning technology, grounded in the cognitive psychology of learning and memory, which is designed to improve learning from and engagement with large-class lectures. The technology allows students to complete weekly multiple-choice quizzes in lectures, submit their answers in a text message using their own mobile phones, and receive immediate feedback. It uses gamification to make the quizzes more engaging: Students track their performance from week to week on a leader board, and earn virtual trophies for achieving certain goals. I describe the design of the system (based on the architecture covered in Reimers & Stewart, 2009), the implementation of a prototype in two 10-week lecture courses, and the results of a survey examining student perceptions of the system. I will also describe the future potential of the system.

- 11:30 Dylan Blaum (Northern Illinois University), Simon Hughes, Peter Hastings (DePaul University), M. Anne Britt, Patty Wallace (Northern Illinois University), Project READi (University of Illinois, Chicago), *Classifying students' causal explanation quality and detecting improvement*

Building on recent success in automatically inferring the quality of written explanations (Hughes et al., 2015), we are examining methods of detecting within-student improvements in essay quality. 1,000 high-school students completed two reading-to-write activities to learn about the causes of two scientific phenomena from multiple documents (before and after an intervention). Humans evaluated the essays for four hierarchical levels of quality: (1) No core content (irrelevant or vague), (2) No causal chains (mentioned elements without connections), (3) Causal chain(s) without intervening (elements directly linked to the outcome), (4) Chain(s) with intervening (successfully included intervening elements). Inter-rater reliability for human scoring



was high (Kappas of .87 and .93). Automatic scoring used window-based tagging models trained to label each word with an element code, one model per code. The maximum and minimum probabilities assigned by each model per word were computed across each entire sentence, and fed into another logistic regression model using a form of 'stacked generalization'. This model was trained to predict the elements and causal relations at the sentence level. Heuristics were used to compute metrics from these predictions to assess the quality. Human and automatic scoring for quality were highly correlated for both topics ( $r = .64$ ,  $r = .59$ ) for pretest performance. We are currently evaluating automated techniques to detect change within an individual that can then be used to provide feedback to students (e.g., attend to directional modifiers, encourage chaining). We will describe the human scoring and the automatic techniques and discuss their strengths and weaknesses.

- 11:45 Elizabeth M. Cedillos-Whynott, Christopher. R. Wolfe (Miami University), *Assessing the effects of video-games on the development of problem solving and creative thinking*

"Don't Starve" is a sandbox, survival, action-adventure video game with the goals of collecting, crafting, and surviving varying elements within the game environment. I present methods for assessing the effects Don't Starve and similar video games have on children's problem solving and creative thinking. Elements within the game can be altered to include more or less of certain items such as food and enemy monsters. Manipulating the frequency with which players encounter elements and items within the environment is hypothesized to affect game-play. By manipulating game-play I hypothesize that we can manipulate the types of learning developed in younger children. More specifically, spending weeks playing a more constructive version of the game (i.e. building tools and structures) may build problem solving skills used to overcome obstacles within the game, whereas, playing a version that is more exploratory, involving exploring the environment and searching for resources, requires a more adaptive type of skill such as divergent or creative thinking. Using the video capture software, Fraps, real-time video-game play can be unobtrusively recorded and various game-play actions can be monitored for trends associated with constructive and exploratory play. I present methods for measuring creativity in constructing tools and structures, as well as overcoming problems and obstacles within the game. Finally, I present methods to determine whether problem solving and divergent thinking skills are transferable to real-world situations outside the video game. Preliminary data from experiments employing this approach are presented to evaluate the viability of these methods.

- 12:00 Aiping Xiong, Weining Yang, Ninghui Li, Robert W. Proctor (Purdue University), *Improving detection of phishing attacks by directing users' attention to domain highlighted URLs*

We conducted an MTurk online user study to assess whether directing users to attend to the address bar and the use of domain highlighting lead to better performance at detecting fraudulent webpages. 160 participants were recruited to evaluate the trustworthiness of webpages (half authentic and half fraudulent) in two blocks. In the first block, participants were instructed to judge the webpage's legitimacy by any information on the page. In the second block, they were directed specifically to look at the address bar. Webpages with domain highlighting were presented in the first block for half of the participants and in the second block for the remaining participants. Results showed that the participants could differentiate the legitimate and fraudulent webpages to a significant extent. When participants were directed to look at the address bar, correct decisions increased for fraudulent webpages ("unsafe") but decreased for authentic webpages ("safe"), suggesting a shift in bias toward fewer acceptances of all webpages, overall. The percentage of correct judgments showed no influence of domain highlighting when the decisions were based on any information on the webpage, but

highlighting increased the percentage correct when participants were directed to the address bar. Thus, domain highlighting, as used in contemporary browsers, was helpful for detecting phishing webpages, but only when participants were told to look at the address bar. This result suggests that a precondition for domain highlighting to be successful in everyday Internet browsing is a method for getting users to attend to the address bar.

**12:15-14:00 Lunch & Posters (Stevens Salon D)**

- 1 David A. Washburn (Georgia State University), *A comparison of methods for analyzing individual differences in small-N designs*

Tremendous progress has been made in understanding the relation between cognitive constructs (e.g., working memory capacity and fluid intelligence) using the psychometric approach, in which intertask correlations reflecting individual differences between participants reveal overlap and dissociations between processes. Comparative studies typically have too few animals to employ these paradigms, and yet robust individual differences between animals have been reported that seem likely to be reliable and informative. Fortunately, for many studies in comparative cognition small sample sizes are offset by large sampling rates, so that each animal may produce thousands of trials on each task. In the present study, a comparison is reported of various strategies for handling individual-difference data, using responses of four rhesus monkeys on a battery of computer-based cognitive tasks. The latent factors and between-variable relations revealed in this task battery will be presented, and implications of the assumption (and potential violations from these assumptions) of the factor analytic and modeling approaches will be discussed. The results highlight both the similarities and the differences across monkeys, and between monkeys and a large-N sample of human participants tested on the same computerized battery.

- 2 Youxuan "Lucy" Jiang, Christopher R. Wolfe (Miami University), *Improving education about women's heart health with a web-based intelligent tutoring system*

Heart disease is the leading cause of death for both men and women in the U.S. Although more women die of heart disease than men each year, women's risk has been underestimated both in and outside the medical community due to the misconception of heart disease as a "man's disease." The present work describes a web-based intelligent tutoring system, Avatar HD, which help users develop gist comprehension of heart disease in women by applying fuzzy-trace theory (FTT). It engages users in one-to-one tutoring dialogues by using a talking avatar and simulates teacher-student interaction by answering user-input questions about heart disease from a linked database, where questions addressing the same topic are linked to the same answer. The purpose of this work is to create an accessible, effective, and engaging online education mechanism to dispel the misconception of heart disease as a man's disease and improve understanding of heart disease in women. The Avatar HD system was assessed using 90 healthy undergraduate students and compared to an educational website with information adapted from the Centers for Disease Control (CDC) website. Results indicated that Avatar HD significantly improved users' declarative knowledge, risk assessment ability, and recognition of signs of heart disease and gender stereotypes unique to women. Results also revealed that Avatar HD significantly increased users' willingness to communicate with others and learn more about heart disease more than the CDC website, which provides a new perspective in heart disease education targeting women.

- 3 Qinyu Cheng, Yehui Liu, Haiying Li, Zhiqiang Cai, Arthur C. Graesser (University of Memphis), *Session recovery of AutoTutor system with rich media client*

AutoTutor is an intelligent tutoring system (ITS) integrated with conversations, animated agents, and tutoring technologies. In AutoTutor, the conversation involves one human learner with one agent (dialogs), two agents (trialogs), or multi-party agents. However, during the online lessons, the lesson is interrupted by various issues, such as unstable Internet or incorrect operation. Therefore, the function of instant system recovery seems substantially critical for online course in the ITS environments. The immediate, automatic system recovers the online lesson with the adaptation of the expected time periods so that the user could continue the online lesson without restarting the system. In addition, the system keeps a complete, neat, and accurate user log file for further data analysis without redundant information. The system is much more adaptive, flexible, and friendly. The implementation of system recovery consists of four steps: (1) User's Information. The system creates a session for recording the user's information (e.g., ID, name) and login time; (2) Log File. The log file automatically records which lesson and which media file the student is working on. Moreover, the log file records the user's action (e.g., clicking, typing, response time); (3) Interruption Detection. When the lesson is interrupted, the system examines whether the user session is valid. If it is valid, the system requires the user to choose CONTINUE or RESTART; (4) Recover. If the user selects CONTINUE, the system reloads the related animated agents' conversation and media. Subsequently, the system displays the media that is interrupted previously.

- 4 Dawn Carroll (California State University, Fullerton), *Technology dependency: Social media and language behavior*

In today's modern society, people have become slaves to technology. The individuals must check "Facebook" or texts throughout the day. Our interpersonal skills have shifted towards people becoming more dependent on technology to start a conversation, rather than speak face to face. However, more than interpersonal skills are being affected by this influx of technological dependency. Our ability to write proper sentences has been affected as well. In texting and social media speak, we no longer use proper punctuation. Instead, our punctuation is substituted by emojis and acronyms. A person might say "lol" or "haha" out of habit to end a sentence, regardless if the sentence calls for it or not. Emojis are used similarly. Emojis are the way people incorporate emotion into their sentence. Without emoji, people would misinterpret what a person is saying due to lack of emotion. "Lol" and "haha" are also ways people attempt to incorporate emotion into a message. However, as previously stated, people use certain emojis, as well as acronyms out of context. This goes to show that our previously programmed knowledge of proper grammar is still being incorporated into the messages in the form of mock punctuation. I will utilize Google Docs to create surveys on how people use punctuation in social media and texting speech. In these surveys, I will utilize multiple choice format, Likert scales, and text examples. I will also utilize the website "Texts From Last Night" to determine which U.S. state has the most common punctuation substitutions.

- 5 Jon D. Elhai (University of Toledo), Brian J. Hall (University of Macau), *Anxiety about data breaches: Results from a community sample*

Electronic data breaches and internet "hacking" are common threats to individuals' security, frequently reported in the news media. Despite public concern about data compromises, little research has examined associated anxiety and stress. Our aim was to determine the anxiety, worry and stress associated with possible internet hacking. We utilized an online survey in fall 2014, sampling 304 adult community participants from North America, recruited through

Amazon's Mechanical Turk labor marketplace. We assessed resting anxiety, and anxiety and stress related to several different forms of data breaches (e.g., hacked email, hacked social media account, etc.). Anxiety ratings from these types of data breaches fit well as a single, underlying factor using confirmatory factor analysis. Ratings of data breach anxiety were greater than resting anxiety. Higher resting anxiety was not related to data breach-related anxiety, indicating that data breach anxiety is specific. Most demographic characteristics were unrelated to data breach anxiety. Anxiety related to data breaches is important to study. Education and intervention are needed to reduce anxiety and increase internet security.

- 6 Jon D. Elhai, Jason C. Levine (University of Toledo), Brian J. Hall (University of Macau), *Anxiety about electronic data breaches and relations with digital privacy protection*

Many recent incidents of hacked digital accounts, or data breaches, have been featured in the news media. Less is known about emotional reactions from the prospect of such breaches. We sampled 305 community participants in spring 2015 from Amazon's Mechanical Turk labor platform. We surveyed subjects using standardized measures of recent anxiety, anxiety from data breaches, and perceived stress. We also inquired about issues of digital privacy, including news exposure, perceived privacy importance, privacy self-efficacy and protection behaviors, and previous data breach victimization. We found that data breach anxiety and recent anxiety were correlated with perceived stress, but differentially. Using multi-group invariance testing, profession (information technology-related or not) moderated the symptom structure of recent anxiety, but not data breach anxiety. Using structural equation modeling, data breach anxiety was related to prior hacking victimization (positively). Digital privacy protection behaviors were related to data breach anxiety and privacy self-efficacy. Data breach anxiety mediated relations between prior victimization and privacy protection behaviors. Privacy self-efficacy mediated relations between news exposure to data breaches and privacy protection. Results help explain determinants and implicate intervention for anxiety related to data breaches.

- 7 Adrian Cunningham, Curt Burgess (University of California, Riverside), *Life beyond semantic distance using HAL and HiDEx*

HAL (Hyperspace Analogue to Language) is a computational high-dimensional semantic model that learns word similarity by "reading" a corpus of language. HiDEx (The High Dimensional Explorer) is software that can process a corpus of text and create word vectors that contain co-occurrence information for all the words in the corpus. It allows the user to build and analyze many variations of the HAL model (Shaoul and Westbury, 2010). Comparisons between these 2 models have been previously discussed (Shaoul and Westbury, 2006, 2010, 2011). These papers were successful in shedding light on many of the differences between HAL and HiDEx. There are, however, more language metrics that can be generated from the HAL model than semantic distance and density. We have two goals in this paper. First, we will make an exhaustive feature comparison between the models and discuss the cognitive and psycholinguistic relevance of the features. Second, we demonstrate the use of the more unknown metrics of the HAL model by using a semantic priming methodology and lexical decision and naming tasks. In addition to semantic distance and density, we will discuss the various Minkowski metrics and their different applications, conditional probability, vector density, as well as neighborhood manipulations.

- 8 Curt Burgess (University of California, Riverside), Justin Estep (San Diego Mesa College), Sarah Maples (University of California, Riverside), *Semantic and grammatical characteristics of closely related verb types*

High-dimensional memory models (e.g., HAL, LSA, Beagle, HiDEx) are best known as semantic models. Less known is that the HAL model (and presumably the others?) has been used to demonstrate a number of grammatical and syntactic effects. The model's meaning representations are sufficient for categorizing nouns, verbs prepositions and determiners (Burgess & Lund, 1997). In that paper, it was also shown that the model was able to distinguish morphological ambiguous and unambiguous verbs (ie, past-tense: knew, drove, past-participle: seen, grown, and those that are morphologically ambiguous: listened, watched). A neighborhood analysis reflects the semantically subtle distinction between "praise" and "flatter." In this analysis, we see if a subtle grammatical/semantic distinction can be detected with the model. We chose two pairs of verb types that illustrate the nature of the physical contact explicit in the word (poke-touch verbs and hit-swat verbs; see Levin, 1993). An ANOVA comparing the intra-group distances (distances of "poke" verbs to other "poke" verbs and "touch" verbs to other "touch" verbs) and extra-group distances (distances of "poke" verbs to "touch" verbs). If intra-group distances are reliably shorter than extra-group distances, one can conclude that the classes are contextually independent. The same analysis was conducted for the hit-swat verbs. The results will be discussed with respect to both the semantic characteristics of the verbs and the thematic nature of how the verbs are typically used in sentences.

- 9 Benjamin Schloss, Ping Li (The Pennsylvania State University), *Predicting human brain activation for concrete nouns: Comparing vector space and graph theory models*

Computational models of semantic representation like Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), Hyperspace Analogue to Language (HAL; Lund & Burgess, 1996), and others (Li, Farkas, & MacWhiney, 2004; Zhao, Li, & Kohonen, 2011) have been important in providing a computational account of how semantic concepts are both formed and represented. However, work connecting these computational models to neuroimaging data remains relatively scarce. Work on the biological plausibility of these models was jumped start by a study that used corpus-driven computational models to predict brain activation for concrete nouns (Mitchell et al., 2008). Additional studies followed up and extended this approach (for a review see; Jiang, Yu, & Huang, 2013). Using publicly available fMRI and computational data from Mitchell et al. (2008) and Fyshe, Talukdar, Murphy, & Mitchell (2013), we present findings that further test these models' interpretability and plausibility. In particular, we compared semantic vector space models like LSA and HAL (Fyshe, Talukdar, Murphy, & Mitchell, 2013), showing that they perform competitively when compared to the original model of Mitchell (2008). As found with semantic space models, we identified a dimensionality effect, showing that 300 dimensions seem to be optimal not just for simulating behavioral data but also for predicting fMRI data. We also compared the prediction performance of vectors that were systematically varied in terms of their window size, whether the underlying frequency counts included phrases or only single words, and whether single word counts were controlled for part of speech or not. We found that vectors that used smaller windows like HAL (as opposed to whole document windows like LSA) and that only counted instances of a target word if it was a noun (controlled for part of speech) performed the best. Finally, we conducted exploratory analyses using graph theory to extract higher quality vectors from smaller text samples and compared these to current methods. This may be particularly useful for researchers studying special and clinical populations who are not represented by existing large text corpora. We discuss our findings and the limitations of our research in regard to the larger field that examines the relationship between computational and

neural representations of meaning. The scripts and programs for our analyses will be made available to the larger research community.

- 10 Melissa Ray, Joe Magliano (Northern Illinois University), *The automatic assessment of moment-to-moment comprehension processes*

This study explored the extent to which adult readers' moment-to-moment processing was influenced by changes in text structure. The Reading Strategy Assessment Tool (RSAT) was used to measure the extent that participants engaged in inference processes associated with constructing a mental model for texts: bridging and elaborative inference. RSAT automatically codes typed "think aloud" protocols for evidence that these processes occurred. In order to identify changes in text structure, a content analysis was to identify the different kinds of relationships that can exist between sentences (e.g., causal, descriptive, comparative). In study 1, participants read informational texts (science and history) presented in RSAT and produced typed "think aloud" protocols after each sentence. In Study 2, participants read the texts silently. Hierarchical linear regression analysis indicated that structural relationships predicted reading times. Causal relationships were correlated with a decrease in reading times, whereas descriptive relationships were correlated with an increase in reading time. Additionally, analyses indicated that while structural relationships were positively associated with bridging inferences, there was some evidence of a negative association between structural relationships and elaborative inferences and. Convergences between the RSAT and reading time data suggest that the perception of structural relationships led to bridging inferences, but that some relationships are more effortful to process than others. Whereas in the past, RSAT has been used to capture individual differences in inference process, this is the first study to show it can be used to study moment-to-moment comprehension processes.

- 11 David Allbritton (DePaul University), *Making text analysis accessible with R and Shiny*

Text analysis and text mining are powerful tools for automatically extracting patterns from unstructured texts, such as texts from open-ended survey questions in program evaluation or research. The open-source statistical programming language R includes powerful modules for text analysis such as tm (Text Mining) and NLP/openNLP (Natural Language Processing) (<http://cran.r-project.org/web/views/NaturalLanguageProcessing.html>). Shiny (<http://shiny.rstudio.com/>) is a web application framework for creating interactive web interfaces to data analysis programs in R. Analysis outputs are wrapped in html markup for viewing in a web browser. Input forms allow the user to change analysis parameters interactively and immediately view the results. A Shiny application is presented for exploratory analysis of open-ended text data, such as course evaluation comments. Texts are analyzed based on lexical n-gram frequency, and the results are presented as frequency plots or as word clouds. The user can change analysis parameters (such as minimum frequency, whether to use stemming, n-gram size, etc.) interactively while viewing the results. This tool could be useful for generating coding categories prior to doing a traditional content analysis, or for identifying useful search terms for finding representative examples from a large set of open-ended responses. Future extensions could include the use of cluster analysis and latent semantic analysis to group similar responses. The evaluation of actual course evaluation data is presented as an example of how R and Shiny can make text mining tools accessible to a wider set of (non-technical) users.

- 12 Evgeny Novikov, Irina Vakoliuk, Igor Varchak, Roman Akhupkin, Dariya Shvaiko, Ekaterina Budenkova, Irina Shalaginova (Baltic Federal University Immanuel Kant), *Computer oculography system for central nervous system monitoring using standard digital camcorders*

Some current investigations in physiology and medicine are focused on human analysis of visual information. The process of recording and subsequent analysis of arbitrary and caused eye movements in general is called oculography. In the last decade, in the wake of the development of digital technologies, computer oculography that uses modern methods of mathematical modeling and computer vision is becoming more and more important. Most of the existing computer oculography systems rely on active eye scanning, which is usually a very expensive solution. However, due to development of technologies for generating digital images and general increase in computing power of personal computers and portable devices, methods of passive scanning images are starting to gain popularity. Our method of using standard digital optics can be employed for accurate and reliable localization of eye centers using vector field of image gradients, obtained with frequency of no less than 30 frames per second. Analysis of visualization of changing pupil center position in time (oculogram) allows us to define functional condition of the brain divisions involved in the programming process and regulation of eye movements, in particular the extrapyramidal system. Unlike many others, the proposed method of computer oculography is non-contact, which is especially important when dealing with children and patients with neurological and psychiatric disorders, distant, affordable and does not require special training of the medical staff or researcher.

- 13 Toshio Ohyanagi, Kunihiro Kanaya, Yasuhito Sengoku (Sapporo Medical University), Ada Leung, Lili Liu, Masako Miyazaki (University of Alberta), *Development of a new Bluetooth Low Energy device for measuring accurate reaction time*

Timing accuracy in measuring reaction times (RTs) on computer systems is critical for scientific studies. It, however, is getting worse as hardware becomes faster and software for conducting experiments becomes more complex (Plant 2014, Plant & Quinlan 2013). We have developed a USB device, termed SMART, realized with a Cypress's PSoC1 microcontroller to measure accurate RTs to visual and auditory stimuli (Ohyanagi & Sengoku, 2010). We have been using the device with our new RT tasks to assess inattention of patients in Occupational Therapy practice (Ohyanagi et al., 2014). As Bluetooth Low Energy (BLE) technology is getting popular and most of recent computer systems including tablets such as iPad support the technology, we developed a new version of the SMART device that was designed to incorporate BLE technology, instead of a USB interface, to connect computer systems. The new device, termed SMART BLE, was realized with Cypress's PSoC4 BLE module, and it works with a CR2032 coin cell battery. In this presentation we explain the detail of hardware and firmware design of SMART BLE device and the accuracy of measuring RTs with an iPad. We plan to demonstrate the use of SMART BLE device with an iPad during our presentation. By using the device with an iPad, our patients will be able to conduct the experiments of assessing inattention in their home environment by themselves.

- 14 Josh de Leeuw (Indiana University), *A collaborative, open-source collection of browser-based experiments for teaching demonstrations using jsPsych*

There are a number of websites that offer online versions of psychology experiments for classroom demonstrations. However, these sites typically have at least one of the following shortcomings: (1) creation of new experiments is limited to a small team of individuals, restricting development; (2) running the experiments requires the installation of additional software, restricting usage; (3) source code for the experiments is private, restricting sharing and

improvement. I am developing a website featuring browser-based experiments that addresses these typical problems. The website is contained in a public open-source GitHub project that anyone can contribute to. GitHub's pages feature and built-in rendering engine Jekyll make it possible to add a new experiment to the website by creating a single text file. In addition to classroom demonstrations, instructors can also use the repository as a platform for enabling student-created experiments. Students can fork the repository which creates their own personal copy, allowing them to modify the experiments provided. Experiments on the website are created using jsPsych. jsPsych is a JavaScript library that enables rapid development of browser-based behavioral experiments. Using jsPsych simplifies the process of programming new experiments, allowing instructors and students with little web-based programming experience to create and modify experiments. A final benefit of the project is that the experiments can serve as useful templates and models for the development of research-oriented browser-based experiments using the jsPsych library.

- 15 Janelle Szary, Michael N. Jones (Indiana University), Rick Dale (University of California, Merced), *Model selection using multi-model inference: An example showing interaction dominance in collaborative foraging*

An important part of the human cognitive experience is social: We learn, play, remember, and problem-solve in a context which is often social and sometimes even collaborative. Here, we investigate the degree to which collaboration effects information retrieval processes in dyads. From a complex systems perspective, a system that can be described as the sum of independent processes will have a very different statistical profile than a system in which component parts become coupled or changed through interaction. Different statistical models are used to characterize each of these systems, as well as systems which range from independent to interdependent. We discuss how the multi-model inference technique can be used to fit multiple candidate models to statistical distributions, while accounting for model uncertainty. Specifically, we fit behavioral and computational dyadic foraging data to a range of candidate models (normal, exponential, gamma, lognormal, and Pareto) which are associated with increasing levels of interactivity and interdependence among system components. Our results show that real and simulated dyads both tend to forage in manners consistent with interaction-dependent, interdependent systems, and suggests how behavior shifts during collaboration.

- 16 Melody W. Dye (Indiana University), Daniel Yarlett (Stanford University), Michael Ramscar (Tübingen University), Michael N. Jones (Indiana University), *The temporal dynamics of distributional semantics*

Distributional models of semantics operate on the assumption that the similarity between two words is a function of the overlap between the contexts in which they occur, a principle commonly known as the distributional hypothesis. Prior investigations have shown that, in line with this principle, a word's semantic representation can be manipulated by situating that word in linguistic contexts frequented by a target meaning. Left open to question is the role of temporal dynamics in the distributional construction of meaning. Learning occurs in time, and it can produce asymmetric outcomes depending on the sequence in which information is presented. In particular, discriminative learning models predict that systematically manipulating the semantic context preceding a word should more strongly influence its meaning than should varying what follows. We find strong support for this hypothesis in three experiments that carefully manipulated subjects' contextual experience with words of varying frequencies.



**14:15-15:15 Symposium: Extrapolating human ratings using distributional semantics and machine learning methods: Advances and limitations (Boulevard A)**

Organized by Emmanuel Keuleers (Ghent University), Chris Westbury (University of Alberta)

Human ratings for variables such as age of acquisition (AoA), imageability, concreteness, and affective ratings are widely used in psycholinguistic research. Recent developments in semantic vector models and machine learning have led to a growing interest in the question of it is possible to extrapolate subjective ratings, using unstructured sources of information such as text corpora or mathematical modeling from large datasets of judgments. In this symposium we focus on a few questions. Which subjective ratings can be successfully extrapolated? What are the methods and techniques that lead to successful extrapolation and why? Is correlation between extrapolated ratings and original ratings a sufficient condition to use extrapolated ratings in lieu of original ratings?

14:15 Paweł Mandera, Emmanuel Keuleers (Ghent University), *Predicting item-level effects of relatedness with models based on prediction and counting*

Recently, a new class of distributional semantic models has been proposed (Mikolov et al., 2013) that can be trained on a text corpus and allows to measure semantic similarity between words. In contrast to more traditional models accomplishing similar goals, such as HAL and LSA, the new class of models can be trained iteratively by learning to predict co-occurring words without an explicit representation of the full co-occurrence pattern. We discuss the relevance of the new class of models for psycholinguistic theories and compare them to more traditional models based on a set of tasks involving semantic processing: large dataset of semantic priming (Hutchison et al., 2013), human associations, TOEFL, relatedness and similarity ratings. We conclude that all investigated types of models allow to predict item-level effects of semantic relatedness in behavioral tasks, but the prediction-based models usually offer a better fit to human data. Theoretically, we argue that these models bridge the gap between traditional approaches to distributional semantics and psychologically plausible learning principles. As an aid to researchers, we release pre-computed semantic vectors for English and Dutch for a range of models together with a convenient interface that can be used to extract a great number of semantic similarity measures.

14:30 Chris Westbury, Elena Nicoladis, Jeff Keith (University of Alberta), *Predicting semantics on the fly: Affect-based computational estimates of noun gender judgments in unilingual English speakers*

Human ratings for variables such as age of acquisition (AoA), imageability, concreteness, and affective ratings are widely used in psycholinguistic research. Recent developments in semantic vector models and machine learning have led to a growing interest in the question of whether it is possible to extrapolate subjective ratings, using unstructured sources of information such as text corpora or mathematical modeling from large datasets of judgments. In this symposium we focus on a few questions. Which subjective ratings can be successfully extrapolated? What are the methods and techniques that lead to successful extrapolation and why? Is correlation between extrapolated ratings and original ratings a sufficient condition to use extrapolated ratings in lieu of original ratings?

- 14:45 Geoff Hollis, Chris Westbury (University of Alberta), *A computational modelling approach to understanding human color preferences*

There is a wealth of literature on human color preferences (for reviews, see Whitfield Whiltshire, 1990; Crozier, 1996; Hurlbert & Ling, 2012; Palmer, Schloss, & Sammartino, 2013). However, our understanding of how preferences vary with respect to color properties is very much in its infancy. In part this is because color preferences appear to be context-specific (e.g., Holmes & Buchanan, 1984; Lind, 1993; Schloss, Strauss, & Palmer, 2012) and, more generally, sensitive to interactions between a variety of color properties as well as personal and cultural features (e.g., Guilford & Smith, 1959; Hurlbert & Ling, 2007; Nemcsics, 2009; Schloss & Palmer, 2009; Schloss & Palmer, 2011; Taylor, Schloss, Palmer, & Franklin, 2013; Bakker, Voordt, Vink, Boon, & Bazley, 2015). The end result is that, despite its size, the literature on human color preferences is fragmented. Research on color preferences seems to have waned in the past few decades, or in some cases turned its attention to highly specific, and applied contexts (e.g., Kwaliek, 1996; Saito, 1996; Kiliç, 2011; Baniani & Yamamoto, 2014). We suspect part of the reason for this waning is that theory-driven, experimental methodology is not well-suited for the nature of the problem at hand. Rather, approaches focusing on statistical modeling and machine learning are more suited for understanding the relationship between properties of color appearance, context, and observer judgments. We describe our efforts to statistically model observer judgments of preference for color triads. Practical applications and future directions are discussed.

- 15:00 Jon A. Willits (University of California Riverside), Tim Rubin, Michael N. Jones (Indiana University, Bloomington), *If you want to know which semantic model is best, don't compare semantic models*

Distributional models of meaning are gaining attention both in science and industry as useful tools for simulating and predicting human judgments about language and word meaning. An effect of this increased attention is a proliferation in the number and types of distributional models. Data and behavioral scientists are increasingly interested in making sense of the vast number of semantic models, and understanding which models perform best. In this work, we will argue that comparing full models to one another – such as comparing Latent Semantic Analysis (Landauer & Dumais, 1997) to BEAGLE (Jones & Mewhort, 2007) – is misleading when trying to understand semantic model performance. This is because these models vary in a large number of parameters, and so one model “winning” at a particular task does not help explain which features of the model led to its superior performance. In this work, we argue for taking a more deconstructive approach. Using human judgments of category membership, category typicality, similarity, and semantic relatedness, we show that some semantic models perform better than others. However, we also show that this doesn't lead to any deep understanding of why some models perform well and others do not. We then show that when you look at particular features of the models (such as whether they encode word frequency counts within documents or co-occurrences between words, or how the model manipulates or transforms the semantic space), we can then learn important, generalizable facts about which features of semantic models perform best in certain situations.

**14:15-15:15 Symposium: Methods effects revisited: Current state of the art of personalization, Radex models, seriousness checks, and visual analogue versus Likert scale format in web-based research (Boulevard B)**

Organized by Ulf-Dietrich Reips (University of Konstanz)

SCiP was one of the first conferences to feature research about the methodology of Internet-based research, beginning in the mid 1990s. Over the years, several methods effects were discovered. In the proposed symposium, we revisit four of these, namely personalization, Radex models, seriousness checks, and visual analogue versus Likert scale format and present new evidence that may help to improve these methods and make them more widely known.

14:15 Tim Kuhlmann, Michael Dantlgraber, Ulf-Dietrich Reips (University of Konstanz), *Advantages of visual analogue scales for brief internet-based measurements*

Brief measurements and short questionnaires are a necessity in most research environments, especially on the Internet with voluntary participants. Visual analogue scales (VAS) have shown superior measurement qualities in comparison to traditional Likert-type response scales in previous studies (Funke & Reips, 2012; Reips & Funke, 2008). The present study expands the comparison of response scales to properties of short Internet-based personality scales in a within-design. An online questionnaire measuring Conscientiousness, Sensation Seeking, and Narcissism was filled out by 879 participants. The questionnaire included the seriousness check (Reips, 2002) and contained all constructs in both answer scale versions in a counterbalanced design. Results show slightly improved reliabilities for the VAS versions of the original scales, in comparison to Likert-type scales. This difference increases for short scales of the constructs. To assess the validity of the measurements, age was used as a criterion, as all three constructs have shown non-zero correlations with age in previous research (e.g., Steinberg et al., 2008). Age was assessed directly and via year of birth, showing a high consistency. For Sensation Seeking, the VAS-version of the scale significantly predicted incremental variance in age, controlling for the Likert-type scale in a regression model ( $\Delta R^2 = .024, p < .001$ ). All of the eight single items of the scale showed higher correlations with age than their Likert counterparts. VAS offer psychometric advantages with regard to reliability and validity in Internet-based measurements, especially when using brief scales to assess personality.

14:30 Michael Dantlgraber, Ulf-Dietrich Reips (University of Konstanz), *A new way of factor structure visualization: The Factor Based Radex Model*

Internet based research facilitates large scale data collection resulting in an increased number of variables and cases in the average psychological dataset. Therefore more complex models can be estimated. We argue there is a need for more optimal data visualization of such models. We present a new way of visualizing factor structure using the validation data of two German language intelligence tests (I-S-T 2000 R, Liepmann, Beauducel, Brocke & Amthauer, 2007; M-KIT, Dantlgraber, in press). Our visualization model is based on Guttman's (1954) and Marshalek, Lohman and Snow's (1983) classic Radex models, extending them from multidimensional scaling (MDS) to factor analysis. The advantage of the Factor Based Radex Model over common visualizations of factorial models lies in the integration of different factor models within the same figure, facilitating comparisons and the understanding of relationships between lower order and higher order factors. For example, regarding the fit-indices CFI and RMSEA the I-S-T 2000 R should be modelled by three correlated factors and the M-KIT by only one general factor. However, on a general latent level the constructs are not distinguishable in a standard presentation format. The factorial structure is easy to understand when using the Factor

Based Radex Model for its visualization and it improves the understanding of complex data sets from psychometric tests. Application areas, limitations, and possible adaptations of the Factor Based Radex Model are discussed.

- 14:45 Stefan Stieger, Ulf-Dietrich Reips (University of Konstanz) , *Impact of personalization of invitation letters to online studies: Study duration and outcome feedback*

Several personalization methods of invitation letters to online studies are known to raise response rates. For example using the participant's name in the salutation of the invitation letter can raise the willingness to take part in the online study (Joinson & Reips, 2007). In the present online experiment, we analyzed the impact of offering different levels of feedback about the study's outcome (general, individualized, personal score including interpretation) and the assumed duration of the study (5 min, 10 min, 20 minutes; real duration: Modus = 8 minutes, also see Galesic & Bosnjak, 2009). Furthermore, we were interested whether providing too short of a statement of duration in the invitation letter leads to a reactance effect (operationalized as the number of missings, i.e, questionnaire completeness). Although we found a small effect between supposed study duration and non-response (effect size  $r = -.06$ ,  $p < .001$ ; response rates: 16.9%, 16.3%, and 11.6% respectively), announced personalized study outcome feedback had no significant impact on study participation,  $r < .01$ ,  $p = .68$ . Expected study duration had no substantial impact on questionnaire completion ( $r = -.03$ ,  $p = .27$ ) and answering behavior (e.g., lower deviation of answers to the items at the end of the study;  $\eta^2 = .001$ ,  $p = .74$ ) suggesting no reactance effect.

- 15:00 Ulf-Dietrich Reips (University of Konstanz), *Revisiting the seriousness check: A simple method to increase data quality in internet-based research*

With ease of access to online studies, a large number of participants with diverse motivations is taking part. People just browsing the Internet for interesting content may provide useless data when clicking through a questionnaire out of curiosity, rather than providing well thought out answers. Researchers or visitors solely interested in having a look at a study's methodology and materials are frequently forced to submit data even if they are not motivated to provide valid responses. These problems concern all types of publicly accessible Internet-based research. The main problem resulting from the participation of nonserious respondents is the increase of noise that may substantially reduce experimental power. The matter poses a serious threat to the validity of online research. One of the methods to ensure quality of data despite these issues is the seriousness check (Aust, Diedenhofen, Ullrich, & Musch, 2012; Reips 2000). Participants are asked whether they want to seriously participate or not, to only analyze data provided by those who have no other motivations (e.g. looking at the study materials for educational purposes). However, the way the seriousness check is implemented may vary. There are important differences in typical wording between labs and there is a striking empirical difference in the proportion of "nonserious" participants between implementations of the seriousness check, namely at the beginning versus end of a questionnaire. In the present paper I thus report on a 4 (location of seriousness check) x 3 (wording) factorial Internet-based experiment. Preliminary data collected from 165 participants indicate differences in dropout depending on location and wording. Results will be discussed.

**14:15-15:15 Experimental Tools (Boulevard C)**

- 14:15 Stian Reimers (City University London), Neil Stewart (University of Warwick), *Audio presentation accuracy in Adobe Flash and HTML5/JavaScript Web experiments*

We present research that builds on our previous examination of visual presentation accuracy and response time measurement in web-based research (Reimers & Stewart, 2014). Specifically, we examine the feasibility of using auditory stimuli in Flash or HTML5 web-based studies. We first discuss different methods with which audio can be presented online. Second, we use a black box toolkit to examine the accuracy of durations for auditory stimuli, and synchronisation of auditory and visual stimuli, with different methods of auditory presentation. Finally, we look at the variability in auditory performance across a number of different computer systems, varying in hardware and software configuration, and using different methods for presenting audio.

- 14:30 Felix Henninger (University of Koblenz-Landau, Max Planck Institute for Research on Collective Goods, University of Mannheim), Pascal J. Kieslich (University of Mannheim), *Psynteract: A flexible, cross-platform, open solution for interactive experiments*

We introduce a novel platform for interactive studies, that is, any form of study in which participants' experiences depend not only on their own responses, but also on those of other participants who complete the same study in parallel. The software thus especially serves the rapidly growing field of social dilemma research within psychology as well as experimental and behavioural economics. In contrast to all available software packages, our platform does not handle stimulus display and response collection itself. Instead, we provide a mechanism to extend existing experimental software to incorporate interactive functionality. This approach allows us to draw upon the capabilities already available, such as accuracy of temporal measurement, integration with auxiliary hardware such as eye-trackers or (neuro)physiological apparatus, and recent advances in experimental software, for example capturing response dynamics through mouse-tracking. In addition, by using the same communication mechanism across software packages, we also enable interoperability between systems. Our source code, which provides support for all major operating systems and several popular experimental packages, can be freely used and distributed under an open source license. The communication protocols underlying its functionality are also well documented and easily adapted to further platforms.

- 14:45 Sébastien Lê, Tâm Minh Lê (Agrocampus Ouest), *Holos: A collaborative environment for holistic approaches*

The aim of this presentation is to introduce a new collaborative environment, called Holos, in which researchers can carry out experiments based on holistic approaches, and then share study resources (stimuli and data collected), either partially or totally, within the scientific community. Holos is the combination of a software for collecting data and a database for storing study resources and their respective results. The software is partially inspired by TCL-LabX (Gaillard, 2009), a free interface dedicated to Sorting for psychoacoustic experiments, except that it deals with a wider range of stimuli (texts, images, sounds, and videos), and a wider range of methods including Sorting and Napping. This software is an android application with a tactile human machine interface in which subjects can easily conduct experiments using a tablet. Stimuli are displayed as icons that can be dragged with one finger, depending on the way they are perceived. When required, subjects can write down information to describe the icons. As subjects perform the task, the database records their finger movements; more

precisely, the trajectories of the icons when they are dragged: we have called this new concept digit-tracking, with reference to eye-tracking. The data recorded enable to study over time the evolution of the dimensions that structure the way the stimuli are perceived individually. In this presentation, we will present different applications of Holos, as for instance, how ten luxury fragrances can be compared, from a sensory point of view and from an advertising point of view, on the other hand.

15:00 Bonnie Perdue (Agnes Scott College), *Computers and comparative cognition research at the zoo*

Comparative cognition research often utilizes computerized technology to assess cognitive abilities in various species. Laboratories have been the most common setting for this type of research historically and at present. However, more recently, researchers at zoos have begun to implement computerized test apparatus to provide enrichment to animals and gain insight into the cognitive abilities of a much wider range of species. Here I will present an overview of some of the research involving computers at the zoo. This includes efforts geared directly towards animals as well as computerized devices that animals might use on exhibit for the visiting public to observe. The focus will be an ongoing project training sun bears to use a touchscreen computer. Details on the design, training and implementation of this research will be highlighted. The computerized test device has the potential to serve as an important form of enrichment for the sun bears as well as an important source of information on cognition in this understudied species.

### **15:15-15:30 Break 3 (Boulevard Foyer)**

### **15:30-16:30 Presidential symposium: Computational approaches to narrative (Boulevard C)**

Organized by Peter Dixon (University of Alberta)

Narrative is a ubiquitous part of day-to-day life: We read stories, tell each other stories, describe ourselves in terms of stories, and perhaps understand the world in terms of stories. Yet narrative is extraordinarily complex, involving the interplay of language knowledge, social experience, and cultural convention. A computational approach to narrative has the potential to make this complexity manageable and to provide a basis for understanding this uniquely human activity. The speakers in this symposium will discuss computational approaches to narrative from three different perspectives: artificial intelligence (Finlayson), discourse processing (Yeari), and computer gaming (Backe).

15:30 Mark Finlayson (Florida International University), *Computing Bartlett: A scientific approach to narrative analysis*

In 1932, F.C. Bartlett, a founder of cognitive psychology, published his landmark book "Remembering." The book described, among other things, experiments involving narratives and storytelling and how they shed light on key psychological and cultural phenomena. By today's standards these experiments were suggestive but not "scientific," and equivalent work since has remained tantalizing outside the reach of precise and reproducible analysis. I explain the fundamental reason why this is so and review recent computational approaches to narrative that is finally bringing Bartlett's ideas (and related work) solidly within the purview of the scientific method. I outline recent results and accomplishments and describe next steps which I believe will revolutionize the study of culture, cognition, and their inter-relationship.

- 15:50 Menahem Yearia (Bar Ilan University), Paul van den Broek (Leiden University), *A computational modeling of semantic knowledge in reading comprehension: Integrating the Landscape Model with the latent semantic analysis*

It is a well-accepted notion that prior semantic (general) knowledge that readers possess plays a central role in reading comprehension. Nevertheless, computational models of reading comprehension have not integrated the simulation of semantic knowledge and dynamic comprehension processes under a unified mathematical algorithm that can be applied on actual behavioral data. The present study introduces a computational model that integrates the Landscape Model of comprehension processes with the Latent Semantic Analysis representation of semantic knowledge. In three sets of simulations of previous behavioral findings, the integrated model successfully simulated the activation and attenuation of predictive and bridging inferences during reading, and the centrality estimation and recall of textual information after reading. Analyses of the computational data revealed new theoretical insights regarding the cognitive mechanisms that underlie the simulated comprehension phenomena.

- 16:10 Hans-Joachim Backe (IT University of Copenhagen), *The glove and the hand: A Lacanian approach to player-avatar relationship*

One of the central challenges in humanist game studies is adaptation of models of world-making, identification, and narrative developed for and established in other disciplines. The computer game appears, simultaneously, as two very different things: a medium with unique communicative properties and a process of human interaction. As game philosopher Miguel Sicart puts it: "we need to consider the game as object, the game as experience, and the process linking both" (Sicart 2011, 64). Theories of the computer game have thus historically moved within a spectrum of blind assimilation of e.g. literary theory without consideration of fundamental differences and exceptionalist thinking that forcefully wanted to re-invent the wheel in order to avoid the resulting errors. While literary theory, media studies, and sociology have been very influential in this process, turning toward psychology for inspiration has been a road less traveled. The presentation will discuss how the Lacanian concept of the mirror stage has been fruitfully applied to the understanding of the relationship between the player and her representation in the gameworld, her avatar, by Bob Rehak (2003). It will then expand on Rehak's by now widely accepted thesis by analyzing a small selection of games in which the relationship is complicated by post-human avatars, arguing that Lacan (by way of Rehak) allows for the discovery of otherwise obscured specific properties of narrative in computer games, namely the meaning-making inherent in the friction between the abilities of the player, the avatar, and the avatar's technologically augmented body.

**16:30-17:30 Business Meeting (Boulevard C)**

## Corresponding Authors

Abney, Drew H., dabney@ucmerced.edu  
Allbritton, David, dallbrit@depaul.edu  
Attar, Nada, nada.attar.cs@gmail.com  
Backe, Hans-Joachim, hanj@itu.dk  
Blaha, Leslie, leslie.blaha@us.af.mil  
Blaum, Dylan, dylblaum@gmail.com  
Burgess, Curt, curt@ucr.edu  
Cai, Zhiqiang, zcai@memphis.edu  
Carroll, Dawn, dcarr010@csu.fullerton.edu  
Cedillos-Whynott, Elizabeth M., cedillem@miamioh.edu  
Cheng, Qinyu, qcheng@memphis.edu  
Conley, Patrick, pconley@uwsp.edu  
Cunningham, Adrian, acunn001@ucr.edu  
D'Mello, Sidney, sdmello@nd.edu  
Dantlgraber, Michael, michael.dantlgraber@web.de  
Dixon, Peter, peter.dixon@ualberta.ca  
Duran, Nicholas, nicholas.duran@asu.edu  
Elhai, Jon D., jon.elhai@gmail.com  
Elhai, Jon, jon.elhai@gmail.com  
Estep, Justin, jestep@sdccd.edu  
Finlayson, Mark, markaf@fiu.edu  
Henninger, Felix, henninger@uni-landau.de  
Howard, Zach, c3145114@uon.edu.au  
Huetten, Stephanie, shuetten@memphis.edu  
Jiang, Youxuan Lucy, jiangy9@miamioh.edu  
Keith, Jeff, jeff.keith@me.com  
Kuhlmann, Tim, tim.kuhlmann@uni-konstanz.de  
Lê, Sébastien, sebastien.le@agrocampus-ouest.fr



Leeuw, Josh de, josh.deleeuw@gmail.com  
Li, Haiying, haiyinglit@gmail.com  
Lippert, Anne, aml237@uky.edu  
Magliano, Joe, jmagliano@niu.edu  
Mandera, Paweł, Pawel.Mandera@ugent.be  
Mueller, Shane T., shanem@mtu.edu  
Novikov, Evgeny, eugen@novikov.de  
Ohyanagi, Toshio, ohyanagi@sapmed.ac.jp  
Paxton, Alexandra, paxton.alexandra@gmail.com  
Perdue, Bonnie, bonnie.m.perdue@gmail.com  
Reimers, Stian, stian.reimers@city.ac.uk  
Reips, Ulf-Dietrich, reips@uni-konstanz.de  
Rubin, Timothy, timrubin@indiana.edu  
Schloss, Benjamin, ben.j.schloss@gmail.com  
Shay, Elizabeth A., eshay@bcs.rochester.edu  
Sheu, Ching-Fan, csheu@mail.ncku.edu.tw  
Singmann, Henrik, singmann@gmail.com  
Stieger, Stefan, stefan.stieger@uni-konstanz.de  
Vinson, David W., dvinson@ucmerced.edu  
Washburn, David A., dwashburn@gsu.edu  
Westbury, Chris, chrisw@ualberta.ca  
Widmer, Colin L., widmercl@miamioh.edu  
Willits, Jon A., jon.willits@gmail.com  
Winston, Patrick, phw@mit.edu  
Xiong, Aiping, xionga@purdue.edu  
Yearia, Menahem, myeari@gmail.com  
Zhao, Xiaowei, zhaox@emmanuel.edu

# PSYCHOLOGY SOFTWARE TOOLS



**Solutions for  
Research, Assessment, and Education**

Providing resources to over 5,000 research institutions and laboratories throughout the world for over 25 years.



Hyperion®



fMRI Digital  
Projection System

Celeritas®



Fiber Optic  
Response System



Ask us about E-Prime training

[www.pstnet.com](http://www.pstnet.com)