

47th Annual Meeting of the Society for Computers in Psychology

Vancouver Convention Center West
November 9th, 2017



Welcome Message from the President



Welcome to the 47th Annual Meeting of the Society for Computers in Psychology (SCiP 2017)! I am very excited about this year's conference. I hope you will agree that we have an extraordinary program for you, which includes 24 submitted talks, one submitted symposium, and 20 posters. We also include two events on the theme of "*Big Data and Deep Learning*." The first is a keynote presentation by Dr. Ping Li from Penn State University, with the title of "*Understanding the Language-Learning Brain with Cyber-Enabled and Computational Methods*". The second is the Presidential Symposium organized by me, with talks by Michael C. Mozer, Kalina Christoff and myself. It is no doubt to me that the keynote and symposium talks will provide an insightful survey of current ideas and approaches on implementing deep learning methods to big data in psychological studies.

I would like to thank Rick Dale, the tireless Secretary/Treasurer of our society. I would not have been able to pull the program together without his help and guidance. I would also like to thank the other three members of the Castellan Award Adjudication Committee: Peter Dixon, Randy Jamieson and Chris Westbury. This year, we have a record-breaking number of student paper submissions for the Castellan Award, and I really appreciate their help.

I hope that you are as excited as me to meet with old friends and make some new ones. If you are new to the conference, please do not hesitate to ask me or members of the Steering Committee questions or to provide your thoughts on the conference.

Thanks for coming to SCiP 2017 and enjoy the conference!

Xiaowei Zhao

Officers of SCiP

President:

Xiaowei Zhao, Emmanuel College

President Elect:

Chris Westbury, University of Alberta

Past President:

Randy Jamieson, University of Manitoba

Secretary-Treasurer:

Rick Dale, University of California, Los Angeles

Federation Representative

Joseph Young

Proceedings Editors

Xiaowei Zhao, Rick Dale

Steering Committee

Alexandra Paxton, UC Berkeley

Brendan Johnson, University at Buffalo

Danielle McNamara, Arizona State University

Stefan Stieger, University of Konstanz

Carol Connor, University of California, Irvine

Ulf-Dietrich Reips, University of Konstanz

Jennifer Trueblood, Vanderbilt University

Matthew Crump, Brooklyn College of CUNY

Stephanie Huetten, The University of Memphis

Castellan Award Adjudication Committee

Peter Dixon

Randy Jamieson

Chris Westbury

Xiaowei Zhao

General information

The Society for Computers in Psychology is a non-profit organization of researchers interested in applications of computers in psychology. Its primary purpose is to "increase and diffuse knowledge of the use of computers in psychological research." Over the past several years the organization has focused on many important issues in psychology, such as computational models of cognitive processes and behavior, computational tools for data collection and analysis, human-computer interaction, knowledge representation in both humans and machines, machine learning, methods and tools for Internet-based research, and technology in the service of improving and evaluating outcomes. We have also encouraged a consideration of cognitively-inspired design of computational technologies and models. Membership is open to any person who has an academic degree and who is active in scientific applications of computers to psychological research.

Location:

Vancouver Convention Centre West (Level 1)

Talks: Meeting rooms 114, 115, 116

Posters: Ballroom B+C (On a side note, this room has floor to ceiling windows with an absolutely spectacular view of the Vancouver Harbor and the mountains, so make sure that you will attend the poster session)

SCiP Time

In order to allow people to visit presentations in different sessions we need to keep a tight schedule. For this reason, we will be timing each presentation according to SCiP time, which is determined by the clock displayed at the registration desk outside of room.

Registration

All attendees of the conference must register. Registration for student is 30 US Dollars or 40 Canadian Dollars and is 60 USD or 80 CAD for Faculty/ Non-student. The cost of registration includes your one-year membership in the society. Onsite registration will be located in the lobby outside of room 115 between 7:30am – 8:30am. You are strongly encouraged to preregister at scip.ws so that we can have you nametag ready.

Program Book

Programs will be available in print at the registration desk and as a PDF at scip.ws. It is recommended that you download the program before come to the Annual Meeting. Internet will NOT be available in the meeting space at Vancouver Convention Centre West.

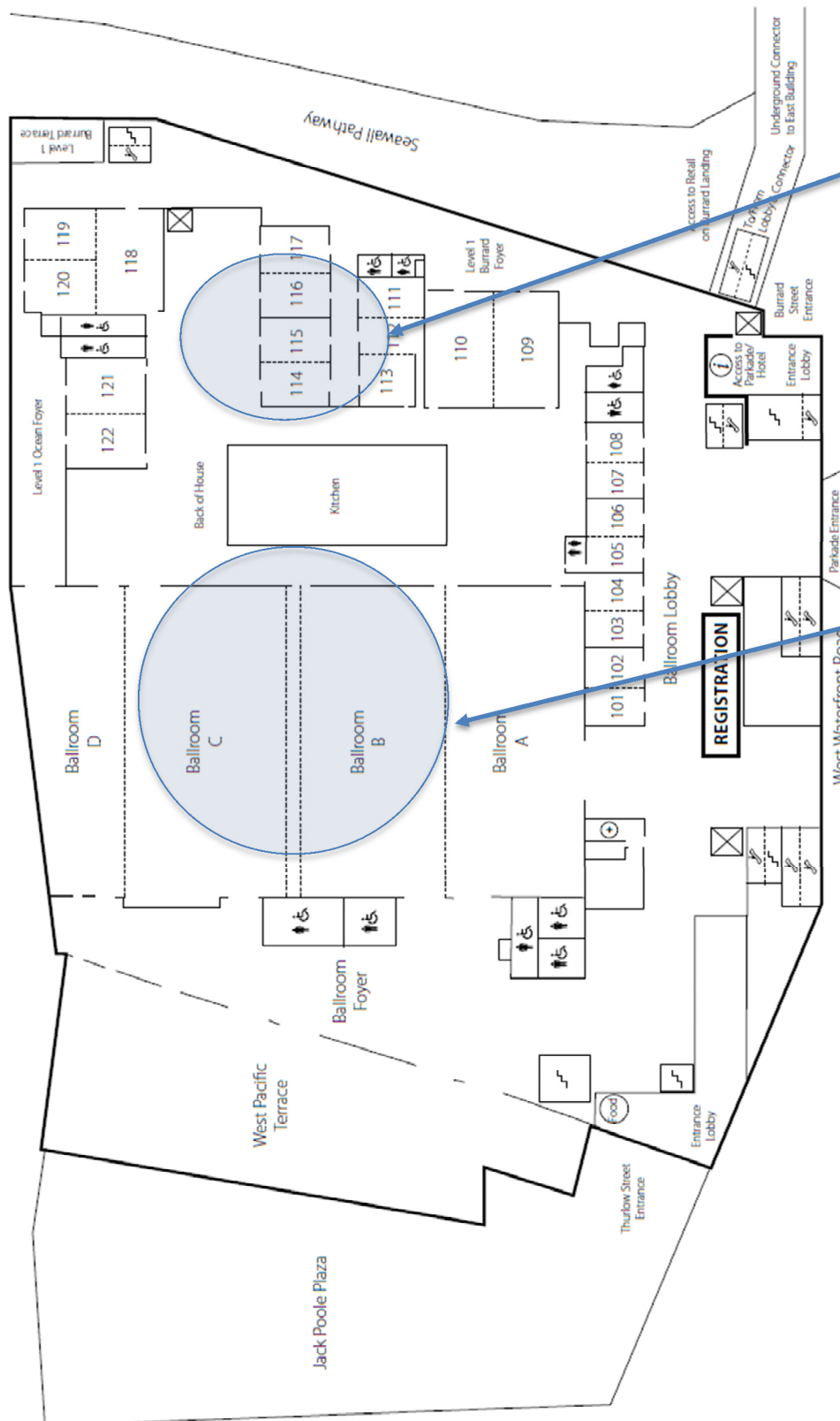
Presentation Guidelines

Digital projectors will be 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 12 minutes, with an extra 3 minutes allowed for questions.

The poster session will be held from 1:00 - 2:00pm at Ballroom B + C. Poster presenters should set up their poster in the period from 12:00pm to 1:00pm and should be available at their poster from 1:00pm to 2:00pm. Posters should be taken down by 2:10pm so conference attendees have time to make their way to the President's Symposium. Posters must fit on one side of a 4' high x 8' wide poster board (could be smaller). Please find your board # on pages 7 and 8.

VANCOUVER CONVENTION CENTRE WEST - LEVEL ONE

Conference Floor Map



Talks (114, 115, and 116)

Posters (Ballrooms B C)

Conference Theme: Big Data and Deep Learning

KEYNOTE

(4:00 pm -5:00 pm, Rooms 114+115)



Dr. Ping Li
Professor of Psychology, Linguistics,
& Information Sciences & Technology
Co-Director, Center for Brain, Behavior, and Cognition
Associate Director, Institute for CyberScience
Pennsylvania State University

*Understanding the Language-Learning Brain with Cyber-Enabled
and Computational Methods*

PRESIDENTIAL SYMPOSIUM

(2:15pm -3:45pm, Rooms 114+115)



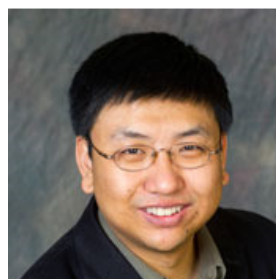
Dr. Michael C. Mozer, Professor
Institute of Cognitive Science and
Department of Computer Science
University of Colorado, Boulder

Predicting Human Learning With Recurrent Neural Networks



Dr. Kalina Christoff
Professor of Psychology
University of British Columbia, Vancouver

*Neural Dynamics of Spontaneous Thought: Evidence from
fMRI and EEG*



Dr. Xiaowei Zhao
President, Society for Computers in Psychology
Associate Professor of Psychology
Emmanuel College, Boston

*Deep Learning of Science Concepts from fMRI data of Naturalistic
Reading*

Time	114		115		116	
7:30 - 8:15	Registration, Refreshments (Lobby outside of the room)					
8:15 - 9:00	Text Analysis (3) Chair: Christopher R. Wolfe		Eye Tracking (3) Chair: Stephanie Huetten			
8:15 - 8:30	Analyzing Tutorial Dialogues About Breast Cancer with Coh-Metrix	Christopher R. Wolfe Colin L. Widmer Savannah J Kreiner Christy V. Torrese & Kennedy E. Braum	Eye Tracking Scanpath Comparisons as a Metric in Naturalistic Visual Search Tasks	Mary Frame, & Richard Warren		
8:30 - 8:45	An Examination of Students' Adaptive Writing Behaviors	Laura K. Allen, Aaron D. Likens, & Danielle S. McNamara	Reading Texts and Comics: Eye Movement Patterns in Adults Using a Coherent/incoherent Paradigm	Lorena A. Martín-Arnal, Ricardo Olmos, José A. León, Inmaculada Escudero, Marcos Ruiz, Manuel Froufe		
8:45 - 9:00	Gender Differences in the Style of Describing Breakup Experiences on Social Media	Lee-Xiang Yang, & Ching-Fan Sheu	Eye-movements during reading in an adult Down's syndrome population: sensitivity to word frequency and general reading patterns	Stephanie Huetten, & Stephanie McMillen		
9:00 - 9:15	Coffee Break					
9:15 - 10:45	Computational semantics (6) Chair: Geoff Hollis		Experimental and Statistical Methods (6) Chair: Pascal J. Kieslich		Internet symposium (6) Organizer: Ulf-Dietrich Reips	
9:15 - 9:30	Using a computational semantic model to quantify form-meaning regularities	Chris Westbury, & Geoff Hollis	Design factors in mouse-tracking: What makes a difference?	Pascal J. Kieslich	<div>1. Personality biases in different types of 'Internet samples' Tom Buchanan</div> <div>2. Perception of time in Internet-based performance tasks: The impact of fluency and response scale Tim Kuhlmann, Michael Dantlgraber, & Ulf-Dietrich Reips</div> <div>3. Alternative scoring procedures and response formats for online achievement tests Birk Diedenhofen, & Jochen Musch</div> <div>4. lab.js: A graphical interface for creating, running and sharing browser-based experiments Felix Henninger, Ulf K. Mertens, Yury Shevchenko, & Benjamin E. Hilbig</div> <div>5. Optimizing animations for Web experiments with high resolution timing requirements Pablo Garaizar & Ulf-Dietrich Reips</div> <div>6. Seriousness check Ulf-Dietrich Reips</div>	
9:30 - 9:45	Collecting Semantic Norms Using Best-Worst Scaling	Geoff Hollis, & Chris Westbury	Improvement of Detecting Spear-phishing Emails by Embedded Skill Training	Aiping Xiong, & Robert W. Proctor		
9:45 - 10:00	Comparing meaning frequency estimates: corpus-based modeling vs. human norms	Barend Beekhuizen, Caitlin Rice, Vladimir Dubrovsky, Suzanne Stevenson, & Blair Armstrong	Computerized Letter Detection Tasks are Not Worth the Effort	Andriana L. Christofalos, & Gary E. Raney		
10:00 - 10:15	Different layers of distributional semantics	Fatemeh Torabi Asr, & Michael Jones	Using video to synchronize audio and visual stimuli in web-based experiments	Stian Reimers		
10:15 - 10:30	Sprout: A Tool for Interactive Topic Modeling and Latent Semantic Space Generation from Small Corpora	Zhiqiang Cai, Art Graesser, Leah Windso, & Xiangen Hu	Survey Automation Detection Methods and its Implications on Psychological Research	John E. Scofield, & Erin M. Buchanan		
10:30 - 10:45	How to Get a Semantic Co-occurrence Model to Distinguish Grammatical Categories at a Fine Grained Level of Representation	Kimberly Miller, & Curt Burgess	TreeBUGS: A User-Friendly Software for Hierarchical Multinomial Processing Tree Modeling	Daniel W. Heck, Nina R. Arnold, & Denis Arnold		

Time	114		115		116	
10:45 - 11:00	Coffee Break					
11:00 - 12:00	Computational/language Modeling (4) Chair: Curt Burgess		Web/Experimental Tools (4) Chair: Alexandra Paxton		Intelligent Tutoring Systems (4) Chair: Danielle S. McNamara	
11:00 - 11:15	Avoiding the Comparing Apples to Oranges Problem in Model Comparison	Curt Burgess, & Katherine Nobel	The Language History Questionnaire (LHQ) 3.0	Fan Zhang, Anya Yu, Xiaowei Zhao, & Ping Li	Let's Walk About That: Random Walk Analyses Contextualize Performance in an Intelligent Tutoring System	Aaron D. Likens, Kathryn S. McCarthy, Laura K. Allen, & Danielle S. McNamara
11:15 - 11:30	Efficient computational modeling of behavioral memory data	Rivka T. Cohen, & Michael J. Kahana	Presenting RDoC app – a web application for psychological assessment of patients	Yury Shevchenko, & Felix Henninger	Stairstepper: Adapting iSTART for Adult Literacy Learners	Kathryn S. McCarthy, Cecile Perret, Aaron D. Likens, Amy M. Johnson, Tricia A. Guerrero, Danielle S. McNamara
11:30 - 11:45	A computational cognitively-inspired technology for clinical diagnosis	Matthew Cook,& Randall Jamieson	Open-source and reproducible workshops: Creating portable learning environments to promote computational education in psychology and cognitive science	Alexandra Paxton, Aaron Culich, Jessica B. Hamrick, Christopher Hench, Todd M. Gureckis, & Thomas L. Griffiths	Recognition of Structure and Meaning of Text in Struggling and Proficient Readers after Interacting with an Intelligent Tutoring System	Anne Lippert, Breya Walker, Raven Davis, Qinyu Cheng, Zhiqiang Cai, Shi Feng, Danielle Clewley, Genghu Shi, & Arthur Graesser
11:45 - 12:00	Integrating Language Modeling and Psycholinguistics	Melody Dye, Michael Jones, & Marti Hearst	Using Inquisit 5 for Online and Mobile Research	Sean C. Draine	Encouraging Deep Question Asking in Tutorial Dialogues Using a Wizard of Oz Technique	Colin L. Widmer, Christopher R. Wolfe, & Christy Torrese
12:00 - 1:00	Lunch and poster setup					
1:00 - 2:00	Poster session (Ballroom BC, Level One)					
2:15 - 3:45	Presidential Symposium (114 + 115) Michael C. Mozer: <i>Predicting Human Learning With Recurrent Neural Networks</i> Kalina Christoff: <i>Neural Dynamics of Spontaneous Thought: Evidence from fMRI and EEG</i> Xiaowei Zhao: <i>Deep Learning of Science Concepts from fMRI data of Naturalistic Reading</i>					
3:45 - 4:00	Coffee Break (Room 116)					
4:00 - 5:00	Keynote Presentation Ping Li <i>Understanding the Language-Learning Brain with Cyber-Enabled and Computational Methods</i> (114 + 115)					
5:00 - 5:30	Business meeting (114 + 115)					

Posters

*** Ballrooms B+C (Level 1) in Vancouver Convention Centre West ***

1 (board # 121)	The time course of semantic and syntactic processing in Chinese word/sentence comprehension: Evidence from eye movements	Xuran Zheng, Lorena A. Martín-Arnal, Ricardo Olmos, & José A. León
2 (board # 122)	Automated Essay Scoring (AES) with Latent Semantic Analysis (LSA) as an alternative to knowledge assessment in front of open-ended questions and multiple-choice tests	José Á. Martínez-Huertas, Olga Jastrzebska, Andrés Martínez, Lorena A. M. Arnal, Ricardo Olmos, & José A. León
3 (board # 123)	Mapping the Theories of Cognitive Science Using LSA	Pablo Andrés Contreras Kallens, & Rick Dale
4 (board # 124)	Complex Communication Dynamics: Exploring the Structure of an Academic Talk	Camila Alviar, Rick Dale, & Alexia Galati
5 (board # 125)	Evaluating methods of keyword extraction from short comprehension passages using word embeddings	Prudhvi Raj Dachapally, Fatemeh Torabi Asr, William R. Aue, Jeffrey D. Karpicke, & Michael N. Jones
6 (board # 126)	Fatherese: Utilizing Corpora to Analyze Male Caregiver Driven Childhood Language Acquisition	Dawn Carroll
7 (board # 127)	EFashiny: A shiny application for exploratory factor analysis	Chi-Lin Yu, & Ching-Fan Sheu
8 (board # 128)	Median Split and Type 1 Error Inflation	Bradley Smith, John Vokey, & Randall Jamieson
9 (board # 129)	Simulating learning in a T-maze with a model of memory: Embodying a psychologically-inspired theory of learning	Bradley Smith, John Vokey, & Randall Jamieson
10 (board # 130)	Applying convolutional neural networks to fMRI data classification: A deep learning approach	Jinlong Hu, Bin Liao, Xiaowei Zhao, & Ping Li,

11 (board # 131)	Speech Emotion detection based on Hilbert-Huang transform (HHT) and deep learning	Zhi-Yong Tao, Ya-Xian Fan, Bo-Jie Zhou, & Xiaowei Zhao
12 (board # 132)	Detecting Contradiction in Agent Source Monitoring During Expository Text Comprehension	Shi Feng, Zhiqiang Cai, & Arthur C. Graesser
13 (board # 133)	The Semantic Cognition in Language (SCiL) Application: Expanding Cognition's Research Frontiers with iOS	Adrian Cunningham, Felipe Ranzani de Luca, Curt Burgess, Bruno Moscardi Flora, Kimberly Miller, Ashima Kumar, Melissa Saidak, Meera Shah, Giselle Urquijo, & Mabel Wong
14 (board # 134)	iPad applications for assessing and training attention function	Toshio Ohyanagi, Kunihiro Kanaya, Yasuhito Sengoku, Lili Liu, Eleni Stroulia, & Masako Miyazaki
15 (board # 135)	Lexical Association and False Memory for Words: A Reexamination Ten Years Later	Yuh-shiow Lee
16 (board # 136)	Does question format change inferences about corrected false information? Comparing elicitation formats in the continued influence effect paradigm	Saoirse Connor Desai, Stian Reimers
17 (board # 137)	Do Constructive Play and Exploratory Representational Play in Video Games Effect Problem Solving and Divergent Thinking?	Elizabeth M. Cedillos-Whynott
18 (board # 138)	Walk and Talk: The impact of attentional load on gait dynamics in young adults	Simon Ho, Kash Daud, Ute Leonards, & Todd C. Handy
19 (board # 139)	Online Survey Tools in Longitudinal Cognitive Psychology Research: A Methodological Review	G. Kyle Gooderham, Simon Ho, & Todd C. Handy
20 (board # 140)	Modeling the moral landscape of political affiliation in the United States	Karie Moorman, & Jeff Yoshimi

Text Analysis

Analyzing Tutorial Dialogues about Breast Cancer with Coh-Metrix

Christopher R. Wolfe, Miami University
Colin L. Widmer, Miami University
Savannah J Kreiner, Miami University
Christy V. Torrese, Miami University
Kennedy E. Braum, Miami University

We describe a method for analyze tutorial dialogues with Coh-Metrix that predicts subsequent comprehension of content about breast cancer. Whereas Coh-Metrix is typically used to make inferences about the mental representations of readers of published texts, we used Coh-Metrix to make inferences about the cognitive processes of interlocutors as dialogue “authors.” We took tutorial dialogues on five related topics from two previous experiments with the BRCA Gist tutor in which questions, pumps, and prompts helped people form “gist explanations.” For both replicates, dialogues were prepared by removing the tutor’s contributions and spell checking, resulting in five paragraph texts for each participant that we analyzed with Coh-Metrix 3.0. As predicted, text Easability Principle Components and Situation Models were highly consistent across both replications. Dialogues scored high on the Easability dimensions referential cohesion and deep cohesion, and low on word concreteness. Dialogues scored relatively high on Situation Model scores for causal particles and causal verbs. After interacting with BRCA Gist, participants completed a Gist Comprehension measure. With respect to Easability, gist comprehension was predicted by word concreteness (both replicates) as well as connectivity and referential cohesion (one replicate). Aspects of the Situation Model predicting gist comprehension include intentional verbs, (both replicates) causal particles, causal LSA, and causal WordNet scores (one replicate). Results contribute to previous finding that tutorial dialogues (specifically coverage of content) are responsible for learning. Results are interpreted in light of theory with a consideration of how key Coh-Metrix functions might be integrated into tutorial dialogues to improve real-time performance.

Corresponding author contact: WolfeCR@MiamiOH.edu

An Examination of Students’ Adaptive Writing Behaviors

Laura K. Allen, Mississippi State University
Aaron D. Likens, Arizona State University
Danielle S. McNamara, Arizona State University

Researchers have recently hypothesized that writers’ abilities to flexibly adapt to varied writing contexts may play an important role in their ability to produce high-quality texts. The current study builds on this work by examining how participants revise texts for different audiences, as well as

whether the properties of these revisions depend on their reading skills. Participants ($n = 95$) were provided with three texts and asked to revise them for two audiences: a group of professors or a class of fourth-grade students. Participants' revised drafts were analyzed along five text Easability components provided by Coh-Metrix (narrativity, syntactic simplicity, word concreteness, referential cohesion, deep cohesion). Results suggest that participants appropriately adapted the texts for the two audiences. Across four dimensions (syntactic simplicity was the exception), participants made the texts easier to read for the fourth-grade students compared to the professors. When participants revised texts for the students, they used language that was more narrative and concrete, and they increased referential and deep cohesion. Importantly, the nature of these revisions depended on the intended audience and participants' reading abilities. On two dimensions (narrativity and referential cohesion), participants with higher reading scores were better able to adapt texts for the audiences than the less-skilled students. This interaction suggests that participants with higher literacy skills engaged in more appropriate revisions at deeper levels of the text. Our results provide important insights into the role of flexibility in writing skill and develop a strong foundation on which to conduct future research and educational interventions.

Corresponding author contact: laura.allen22@gmail.com

Gender Differences in the Style of Describing Breakup Experiences on Social Media

Lee-Xieng Yang National Chengchi University
Ching-Fan Sheu National Cheng-Kung University

A relationship breakup is a stressful event. How people cope with it can often be understood via analyzing their self-reports about breakup experiences. Nowadays, people increasingly share their life events on social media. The breakup stories they post can be regarded as their self-narratives, which could be suitable for research in terms of sample size and anonymity. In this study, we investigated gender differences in coping with breakup via these posts on social media. To this end, we focused on Dcard, a popular social network site among college students in Taiwan. In the data-collecting stage, we first scrapped 36,000 posts on the forum of relationship using R to call the Dcard application program interface (API). Second, posts with breakup in the title ($N = 1,311$) were retained. Among them, about 70% are made by females. No gender difference was found in these posts with respect to the number of likes, comments, or authoring replies to comments. Text mining was performed after word segmentation was done on each post. With the most-frequent words, we found that females mentioned more often about memories, sentiments, and boyfriend's family, whereas males mentioned more often about honesty, change, and commitment. A stylometric analysis was conducted and a difference in writing style between genders was found. Finally, a difference in coping strategy was uncovered by principal component analysis (PCA) in that males tended to emphasize the bright side of past relationship and females were more likely to seek social support.

Corresponding author contact: lxyang@gmail.com

Eye Tracking

Eye Tracking Scanpath Comparisons as a Metric in Naturalistic Visual Search Tasks

Mary Frame, Wright State Research Institute

Richard Warren, Air Force Research Laboratories 711HPW/RHXM

Eye tracking provides a rich source of data in naturalistic surveillance tasks. Understanding the dynamics of eye scanning behavior provides a window into attentional processes of analysts and how this impacts their accuracy and performance over long shifts. Such data can differentiate novices from experts, distinguish scanpaths with varying degrees of fatigue, and assess the relative effectiveness of candidate analyst augmentation aids. However, due to the complexity of the required analysis, the number of data files, and the file sizes, much of the collected and stored data remains unanalyzed or under-analyzed. It can be challenging to make meaningful comparisons using standard eye tracking analysis techniques such as saccade frequency or blink rate in surveillance studies due to the temporal irregularity and occasional overlapping of events of interest. Additionally, data collected from an applied surveillance study can last the duration of a full 8 hour shift, naturally leading to massive eye tracking data files. This can lead to thousands of potential comparisons that need to be made in a single study, yielding output nearly as complex as the eye tracking input data. Solving these problems requires the use of big data techniques. We computed scanpath similarity on a variety of surveillance data sources using algorithms initially developed in making DNA protein sequence comparisons and developed methods of batch processing in Matlab that can parse complex output data on a standard research computer or can be implemented using a supercomputing system for larger or more complex data files.

Corresponding author contact: mary.frame@wright.edu

Reading Texts and Comics: Eye Movement Patterns in Adults Using a Coherent/incoherent Paradigm

Lorena A. Martín-Arnal, Universidad Autónoma de Madrid,

Ricardo Olmos, Universidad Autónoma de Madrid,

José A. León, Universidad Autónoma de Madrid

Inmaculada Escudero, Universidad Nacional de Educación a Distancia,

Marcos Ruiz, Universidad Nacional de Educación a Distancia,

Manuel Froufe, Universidad Autónoma de Madrid,

Theories of visual narrative understanding have often focused on the changes in meaning across a sequence, like shifts in characters, spatial location, and causation, as cues for breaks in the discourse structure (Cohn & Bender, 2017). We examined the relationship between narrative structure and coherence shifts in the segmentation in: a) visual narrative sequences and b) text narrative sequences, in order to compare texts vs comics in eye movement patterns. We used a “coherent/incoherent task” where participants, after viewing three sequences presented in text or comic format, were requested

to select between coherent/incoherent endings of the story (Völlm et al., 2006). Thirty-two undergraduate participants, all of them Spanish native speakers, took part in this study. We used an eye tracker Tobii x-120 to record fixation durations, number of fixations and number of regressions out. For these last ones, results showed that regressions out the coherent ending were higher than the incoherent, and they were also higher for comics than for texts. Besides, in both fixations measures, participants made longer and more fixations in the coherent condition in texts, whereas the inverse results for comics were found only in fixations duration, due to the significant interaction between these two factors. A possible explanation for this effect in text conditions could be that people adopt a searching strategy to find whether the story has an incoherence or not. When viewing comics and reading texts, people may use different processing strategies; despite of that the common goal is the final story comprehension.

Corresponding author contact: loreleimajere@gmail.com

Eye-movements during reading in an adult Down's syndrome population: sensitivity to word frequency and general reading patterns

Stephanie Huetten, University of Memphis
Stephanie McMillen, University of Memphis

Adults with Down's syndrome are able to read, but experience deficits in reading level, speed, and ability to read silently. There are many clinical interventions and remediation strategies, but a great deal of these practices are not based in research investigating which mechanisms involved in reading differ from a typically developing population. The present study uses eye-tracking to compare typically developing undergraduates to an adult Down's syndrome population, examining sensitivity to word frequency. If affected, this may imply a deficit in lexical-level statistical learning mechanisms involved in language learning. Typically developing readers fixate on high frequency words for a shorter duration than low frequency words. The present study replicated sensitivity to word frequency in typically developing undergraduates, and found less sensitivity in the Down's syndrome adults in line with our predictions. Special attention will be paid to eye-tracking methods of testing adults with Down's syndrome. In another exploratory task, participants read a text passage by the paragraph to examine general differences in eye-movements. This more qualitative data set is being used to observe other potential mechanistic differences that influence reading performance and comprehension in Down's syndrome.

Corresponding author contact: shuette@memphis.edu

Computational Semantics

Using a computational semantic model to quantify form-meaning regularities.

Chris Westbury, Department of Psychology, University of Alberta

Geoff Hollis, Department of Psychology, University of Alberta

In their paper on using naïve-discriminant learning to model reading, Baayen et al. (2011) noted that “Morpheme-like effects can be present without an input string requiring a parse into a sequence of morphemes” (p. 56), i.e. words can share both meanings and formal features without sharing a common root. These words are known as phonoaesthemes. A commonly-cited example is the set of words beginning with ‘gl’ that have to do with a short, sharp visual experience: e.g. ‘glitter’, ‘glow’, ‘gleam’, ‘glimpse’. It is predictable that some form-meaning correlation is always going to present for any form, since most words have several morphological neighbors that share both form and meaning. When we select a word by meaning, this form resemblance comes along for free, and vice versa. We present and assess a method for quantifying the degree of form-meaning regularity for any specified form, using a computational model of semantics. We show that the model is able to select words with the same form using the simple vector computations that have traditionally been used to model semantics. The method can also shed light on why phonoaesthetic form-meaning correlations arise.

Corresponding author contact: chrisw@ualberta.ca

Collecting Semantic Norms Using Best-Worst Scaling

Geoff Hollis, Department of Computing Science, University of Alberta

Chris Westbury, Department of Psychology, University of Alberta

Semantic norms have long played a role in verbal learning and psycholinguistic research. Such norms are typically collected using rating scale or numeric estimation formats. We introduce another response format largely foreign to psychology called best-worst scaling and demonstrate its utility for the purpose of collecting semantic norms. In a series of lab-based and crowdsourced studies comparing best-worst scaling, rating scales, and numeric estimation as response formats, we found that best-worst scaling consistently provided higher quality estimates or required less data to arrive at equivalent quality estimates compared to alternative methods. Best-worst scaling presents itself as a promising methodological advancement for both reducing the cost and improving the quality of large-scale, crowdsourced semantic norms datasets that are becoming prevalent within psychology.

Corresponding author contact: hollis@ualberta.ca

Comparing meaning frequency estimates: corpus-based modeling vs. human norms

Barend Beekhuizen, University of Toronto
Caitlin Rice, University of Pittsburgh
Vladimir Dubrovsky, University of Toronto
Suzanne Stevenson, University of Toronto
Blair Armstrong, University of Toronto

Most words are ambiguous, with interpretation dependent on context. Increased understanding of the disambiguation process can be achieved by developing computational models of word meaning based on properties of their linguistic context, and comparing these models to psycholinguistic experiments. We evaluated several word sense disambiguation models on homonyms in natural language corpora. Homonyms are particularly important to understand given their prevalent use in psycholinguistics, with the clear cut distinction between meanings offering a foothold for evaluating models before tackling fine-grained ambiguities. We compared model performance against behavioural norms of relative frequency of each homonym's meaning from [Armstrong et al. 2012; BRM 44(3), 1015-1027]. Each model was used to identify the meaning of all occurrences of 553 homonyms in everyday corpora (Wikipedia, video subtitles). We correlated the relative meaning frequencies for each homonym with the norms and with performance in lexical decision. We observed significant or numerical correlations between the models and human data, highlighting the promise of these approaches. However, the models' estimates of dominant meaning frequency were typically higher than the norms. Manual inspection of a sample of annotations indicated that they were generally correct. Thus, the discrepancy may not be fully attributable to poor model performance. Furthermore, observing this pattern in estimates derived from the same corpora used to derive other psycholinguistic measures suggests that these corpora are representative of meaning usage. These results raise new questions regarding what factors human relative meaning frequency estimates are sensitive to, and how to incorporate those factors into models of ambiguity resolution.

Corresponding author contact: barendbeekhuizen@gmail.com

Different layers of distributional semantics

Fatemeh Torabi Asr, Indiana University
Michael Jones, Indiana University

Word embeddings obtained from neural networks trained on big text corpora have become popular representations of word meaning in computational linguistics. The most popular model recently, i.e., word2vec, simultaneously generates a set of word and context embeddings, the latter usually discarded after training. We demonstrate how these two layers of distributional representation can be used in predicting taxonomic similarity vs. asymmetric association between words. Our study is composed of both artificial language experiments and evaluations based on word similarity and relatedness datasets collected through crowdsourcing and psycholinguistic experiments. In particular, we use two recently published datasets: SimLex-999 (Hills et al. 2016) including explicitly instructed ratings for word similarity, and explicitly instructed production norms (Jouravlev & McRae, 2016) for word relatedness. We find that people respond with words closer to the cue within the context embedding space (rather than the word embedding space) when they are explicitly asked to generate

thematically related words. Taxonomic similarity ratings are however better predicted by word embeddings alone. This suggests that the distributional information encoded in different layers of the neural network reflect different aspects of word meaning. Our experiments also elaborate on word2vec as a model of human lexical memory by showing that both types of semantic relations among words are encoded within a unified network through reinforcement learning. Recommendations for biasing the model to organize words based either on taxonomic similarity vs. relatedness are introduced for practical applications.

Corresponding author contact: torabiasr@gmail.com

Sprout: A Tool for Interactive Topic Modeling and Latent Semantic Space Generation from Small Corpora

Zhiqiang Cai, University of Memphis
Art Graesser, University of Memphis
Leah Windsor, University of Memphis
Xiangen Hu, University of Memphis

Topic modeling and latent semantic analysis (LSA) have been playing important roles in text analysis. Topic modeling provides interpretable topics in a target corpus and LSA provides a way to represent meanings of texts. However, researchers have been frustrated with non-interpretable topics and unacceptable semantic representations when the target corpus is not very well composed for topic modeling and LSA space generation. The documents in a corpus could be too short, such as tweets, conversational utterances. A corpus might be too specific and does not contain enough language support. Sprout is a tool we created to solve these problems. Sprout allows researchers to start with a relatively small set of short documents, called “seeds”. The tool automatically identifies key terms in a target corpus and expands the corpus by adding semantically relevant articles from specified resources, such as Wikipedia. Topic modeling and LSA space generation can then be run on the expanded corpus, which provides more meaningful topics and semantic representations. Sprout also provides interactive topic modeling. Sprout displays top words in each iteration of the modeling. Researchers can put in “seed” topics at any time by typing in a line of words to direct a topic to “grow” around the seed words. Sprout has a visualization module that projects topics and documents contained in a target corpus to a plan. Topic intrusion and LSA nearest neighbors are provided in the tool for immediate validation of topics and semantic representations.

Corresponding author contact: zca@memphis.edu

How to Get a Semantic Co-occurrence Model to Distinguish Grammatical Categories at a Fine Grained Level of Representation.

Kimberly Miller, University of California, Riverside
Curt Burgess, University of California, Riverside

Global co-occurrence memory models such as HAL and LSA, as well as more recent versions of co-occurrence models, are typically described as semantic in nature. Evidence for this characterization includes a wide range of phenomena including, among others, semantic priming and categorization,

document sorting, and scene recognition based on object meaning. However, our earlier research with the HAL model has shown that grammatical sensitivity is reflected in the meaning vectors as well. The model is able to distinguish between the most basic grammatical categories such as verbs, nouns, determiners, and adjectives. The model can also make a more finer-grained distinction between manner of motion verbs and inherently directly verbs and past-tense and past-participle verbs. However, we have not been successful with demonstrating that the model can make very fine grain distinctions such as the get and obtain subclasses of give verbs. Following Levin's (1993) theory, we evaluated HAL's ability to make these fine-grained distinction and report success. Our breakthrough hinged on discovering that transforming the raw weighted co-occurrence values in the word representations into a form that has a history of better predicting human reaction times. These results will be discussed in the context of the behavioral regularities of core meaning and argument structure and Levin's theory of verb relationships.

Corresponding author contact: curt@ucr.edu

9:15am – 10:45am

Room 115

Experimental and Statistical Methods

Design factors in mouse-tracking: What makes a difference?

Pascal J. Kieslich, University of Mannheim

The analysis of mouse movements to study cognitive processes is becoming increasingly popular in many psychological research areas. When creating mouse-tracking experiments, researchers face a number of design choices, for example, about the mouse speed, starting procedure, or response indication. In previous research, many different settings have been employed, but so far little is known about how these methodological differences affect mouse-tracking data. To address this, a series of experiments was conducted to systematically investigate the influence of multiple design factors. A classic mouse-tracking task was used for these experiments, in which participants classify typical and atypical exemplars into one of two categories. In separate experiments, the design factors response indication, mouse sensitivity, aspect ratio, and starting procedure were manipulated between participants. In detail, it was manipulated whether participants could indicate their classification decision by clicking on the corresponding response button or by entering the button area without a click, whether medium or reduced mouse speed was used, whether the screen aspect ratio was 4:3 or 16:10, and how the starting procedure in relation to the stimulus presentation was implemented. Across all experiments and design conditions, the core finding that mouse movements deviate more towards the non-chosen option for atypical than for typical exemplars was replicated. However, the strength of this effect and the distribution of the individual trajectories was influenced by several design factors – especially by the type of starting procedure and response indication. Implications for the design of future mouse-tracking studies and possible recommendations are discussed.

Corresponding author contact: kieslich@psychologie.uni-mannheim.de

Improvement of Detecting Spear-phishing Emails by Embedded Skill Training

Aiping Xiong, Purdue University

Robert W. Proctor, Purdue University

Spam quarantine notification is designed to provide users an opportunity to review all their bulk and junk emails in one place. Instead of wasting time examining legitimate-looking email that actually is not legitimate, users develop the habit of reviewing unsolicited bulk emails and messages carrying malicious payloads or web-links with regular spam quarantine notifications shown in the inbox. Although email spam scanning systems use advanced technology to sort good from bad email, they show deficits in blocking spear-phishing emails which are customized to pass anti-spam filters and targeted to a small number of users. Therefore, while spam quarantine notifications aim to protect users from spam emails, they can also reduce users' suspicion of spear-phishing emails. We conducted a role-play scenario online study to examine users' ability to differentiate spear-phishing emails from genuine emails. Using a between-subject design, we studied four conditions with two factors: spam quarantine notifications and embedded-training of how to identify correct domain names following an error in judgment. One-week later, all participants were asked to manage another set of emails without training. We hypothesized that: 1) The presence of spam quarantine notifications would increase participants' clicks on the links in the phishing emails and their authentication with ID and password to the simulated phishing sites; 2) Participants who get embedded-training would perform better than untrained participants at discriminating phishing and genuine emails. The confirming results imply that, while email spam scanning systems can be improved, a better approach may be to train users' automatized phishing-detection skills.

Corresponding author contact: xionga@purdue.edu

Computerized Letter Detection Tasks are Not Worth the Effort

Andriana L. Christofalos, University of Illinois at Chicago,

Gary E. Raney, University of Illinois at Chicago

During the letter detection task, participants are instructed to read a passage for comprehension on paper while circling all instances of a target letter, such as the letter t. The Missing Letter Effect (MLE) refers to the common finding in which readers miss (do not detect) more target letters in function words (e.g., the, at) than content words (e.g., cat, tuna). We examined whether this well-known effect would be present when using two novel computerized versions of the letter detection task. Participants read easy and difficult passages while performing the letter detection task by circling letters on paper (Experiment 1), by tapping on letters using a touch-screen tablet (Experiment 2), or by clicking on letters using a computer mouse (Experiment 3). The overall pattern of errors was similar across each version of the task, but differences were less consistent in the two computerized tasks. For the paper task, we found a significant MLE and significantly more errors for easy than for difficult passages. We also found a significant interaction, which reflects a smaller MLE for easy than for difficult passages. For the mouse task, we found a significant MLE but no reliable difference between easy and difficult passages overall and no interaction. For the tablet task, we found a non-significant MLE, no reliable difference between easy and difficult passages overall, and no interaction. This

provides an example of when developing computerized tasks is not worth the effort. The various strengths and weaknesses of each methodology will be discussed.

Corresponding author contact: achris29@uic.edu

Using video to synchronize audio and visual stimuli in web-based experiments

Stian Reimers, City, University of London

Recently, we have examined the synchronisation of auditory and visual stimuli in Flash and JavaScript web-based research (Reimers & Stewart, 2016, BRM), and found it unreliable. One alternative to attempting to control and synchronise auditory and visual streams directly is to use video clips containing auditory and visual stimuli. This has the potential for improving synchronisation, albeit at the cost of some flexibility. Using a Black Box Toolkit, I examine the reliability of synchronisation of auditory and visual channels in web-based video stimuli, across browsers and devices, and indicate the circumstances under which it might be productively used.

Corresponding author contact: stian.reimers@city.ac.uk

Survey Automation Detection Methods and its Implications on Psychological Research

John E. Scofield, University of Missouri
Erin M. Buchanan, Missouri State University

Web-based data collection methods such as Amazon's Mechanical Turk (MTurk) are an appealing option to recruit participants quickly and cheaply for psychological research. While concerns regarding data quality have emerged with MTurk, several studies have exhibited that data collected via MTurk are as reliable as traditional college samples and are often more diverse and representative of noncollege populations. The development of participant screening methods, however, has been less explored. Omitting participants based on simple screening methods, such as response time or attention checks may not be adequate identification methods, with an inability to delineate between real or fake participants. An alternative form of suspicious survey responses stem not from human participants, but from survey automation techniques such as survey bots or automated form fillers. The current project develops survey automation detection (SAD) methods while overcoming previous screening limitations. Multiple checks are employed, such as response time, skewness and kurtosis values, and the number of utilized choices from a given range of scale options. This method is tested on a survey taken with an easily available plug-in survey bot, as well as compared to data collected by human participants providing both real and randomized answers. An R function is proposed for researchers to screen for potential problems with MTurk data. Identified cases can then be used as part of sensitivity analyses to warrant exclusion from further analyses. SAD methods can be a promising tool to identify non-real or automated data via MTurk or other online data collection platforms.

Corresponding author contact: John1551@live.missouristate.edu

TreeBUGS: A User-Friendly Software for Hierarchical Multinomial Processing Tree Modeling

Daniel W. Heck, University of Mannheim
Nina R. Arnold, University of Mannheim
Denis Arnold, Institut für Deutsche Sprache

Multinomial processing tree (MPT) models are a class of measurement models that account for categorical data by assuming a finite number of latent cognitive processes. Traditionally, data are aggregated across participants assuming that observations are independent and identically distributed. To explicitly account for participant heterogeneity, hierarchical Bayesian extensions of MPT models instead assume that the individual parameters follow a continuous hierarchical distribution. To facilitate the application of these models, we present the user-friendly and comprehensive R package TreeBUGS, which implements the two most important hierarchical MPT approaches for participant heterogeneity assuming independent beta and multivariate normal distributions of the parameters, respectively. TreeBUGS reads standard MPT model files and obtains Markov chain Monte Carlo samples that approximate the posterior distribution. The functionality and output is tailored to the specific needs of MPT modelers and provides tests for the homogeneity of items and participants, individual and group parameter estimates, fit statistics, within- and between-subject comparisons, as well as goodness-of-fit and summary plots. We also propose and implement novel statistical extensions to include continuous and discrete predictors (either as fixed or random effects) in the latent-trait MPT model.

Corresponding author contact: heck@uni-mannheim.de

9:15am – 10:45am

Room 116

SYMPOSIUM

Current methodological investigations in Internet-based research

Organizer: Ulf-Dietrich Reips, University of Konstanz, Germany

Corresponding author contact: reips@uni.kn

The proposed SCiP symposium will continue the tradition of presenting methods and tools for Internet-based research, with six papers by authors from four countries. Presentations concern the topics of sampling, timing, response formats, study creation, and compliance, specifically *personality biases in different types of 'Internet samples', perception of time in Internet-based performance tasks, high resolution timing in Web experiments, alternative scoring procedures and response formats, a graphical interface for browser-based experiments and seriousness checks*. We present new methodological insights, new tools and new evidence that may help to improve Internet-based research methods and make them more widely known.

Personality biases in different types of 'Internet samples'

Tom Buchanan, University of Westminster

Tom Buchanan (University of Westminster, UK) reports a series of studies exploring whether different methods of participant recruitment lead to personality biases in online samples, and whether such biases could affect research outcomes. His studies (one pseudo-experimental and two correlational designs) show that not only do these biases exist, but indeed may have an impact on research results. Results extend existing work on online data quality, showing that personality biases may arise from recruitment methods.

Perception of time in Internet-based performance tasks: The impact of fluency and response scale

Tim Kuhlmann, University of Konstanz

Michael Dantlgraber, University of Konstanz

Ulf-Dietrich Reips, University of Konstanz

In this presentation, Kuhlmann et al. (University of Konstanz, Germany) report on a 2 (fluency) x 2 (scale format) Web experiment with 571 participants that expands their research on the comparison of visual analogue versus Likert format response scales to *assessments of perceived time in Internet-based performance tasks*. Objective reaction times did not differ between the conditions, but subjective estimates tended to be faster for fluent tasks. In line with previous research measurement with visual analogue scales resulted in better agreement between subjective estimates and objective time differences than with Likert-type scales.

Alternative scoring procedures and response formats for online achievement tests

Birk Diedenhofen, University of Düsseldorf

Jochen Musch, University of Düsseldorf

The third presentation, by Diedenhofen and Musch (University of Düsseldorf, Germany), is also concerned with response format in Internet-based tests and investigates the popular multiple-choice (MC) format that is typically scored by simply counting the number of correctly answered items. In a series of Web experiments the authors examine the MC format in comparison with two alternative response formats: (1) Answer-until-correct (AUC) testing that is suited to capture partial knowledge, and (2) the Pick-N format that extends MC items by implementing more than one correct answer option to impede guessing.

lab.js: A graphical interface for creating, running and sharing browser-based experiments

Felix Henninger, University of Koblenz-Landau
Ulf K. Mertens, University of Heidelberg
Yury Shevchenko, University of Mannheim
Benjamin E. Hilbig, University of Koblenz-Landau

In the fourth presentation, Henniger et al. (from four different German universities) introduce *lab.js*, a free, open-source experiment builder and JavaScript library designed to simplify the realization of browser-based experiments. The tool aims to support researchers with a basic knowledge of HTML and CSS in the construction of studies through a drag-and-drop interface. By adding JavaScript code, studies can be customized further and adapted to complex requirements. In the tradition of HTML- and JavaScript-based Web experiments conducted in the open Web, studies made with *lab.js* can also be shared, archived, re-used and adapted, easily enabling transparent replications, and facilitating cumulative science.

Optimizing animations for Web experiments with high resolution timing requirements

Pablo Garaizar, Universidad de Deusto
Ulf-Dietrich Reips, University of Konstanz

The fifth presentation by Garaizar (Universidad de Deusto, Spain) and Reips, “Optimizing animations for Web experiments with high resolution timing requirements”, continues the topic of Web experiments that use standard web technologies (e.g. JavaScript and HTML) and standard web browsers for research used by Web services like WEXTOR.org and research-oriented frameworks like jsPsych or *lab.js*. The authors review new features provided by the recent revision of web standards in order to optimize animations for Web experiments with high resolution timing requirements and study their accuracy and precision using specialized hardware.

Seriousness check

Ulf-Dietrich Reips, University of Konstanz

One of the methods to ensure quality of data in Internet-based research is the *seriousness check* (Aust, Diedenhofen, Ullrich, & Musch, 2012; Reips 2000). In the symposium’s final presentation Reips looks at effects of different ways of implementing the check. He reports on data from (1) a factorial Internet-based experiment and (2) from a database of 3394 experiments hosted on WEXTOR.org, an experiment generator that has existed since 2000 and implements a seriousness check by default.

Computational/Language Modeling

Avoiding the Comparing Apples to Oranges Problem in Model Comparison

Curt Burgess, University of California, Riverside
Katherine Nobel, University of California, Riverside

In reviewing computational models of semantics, Jones, Willits and Dennis (2015) ask the question “so which model is right?” They argue that models tend to differ with respect to three characteristics: the representational structure of word meaning, the mathematical transformations that transform the co-occurrence information into meaningful representations, or the comparison process used to compute similarity. They further note that the question itself is not very meaningful when comparisons are made using different training corpora, vocabulary lists, and not systematically evaluating the parameters used in the comparison process. In this study, we compared the proficiency of the HAL and LSA models in predicting reaction times to single words, strength of semantic and associative priming and manipulated each model's parameters in this process. Both models were trained on the same ~ billion word text corpus. The results of this process will be reported.

Corresponding author contact: curt@ucr.edu

Efficient computational modeling of behavioral memory data

Rivka T. Cohen, University of Pennsylvania
Michael J. Kahana, University of Pennsylvania

Computational modeling is an important means of describing human memory and identifying its underlying mechanisms. The utility of a model and its associated theory can be assessed via its ability to provide good fits to the observed data, which requires obtaining the parameters that produce the best fits. Efficient algorithms and modeling techniques are needed to obtain a) accurate results in b) a timely manner (i.e., hours vs. days or weeks). This tutorial will provide an introduction to efficient computational modeling techniques for behavioral memory outcomes in free recall tasks. It will begin with a brief overview of the Context Maintenance and Retrieval model (Lohnas et al., 2015), an established framework for modeling behavioral outcomes of free recall memory tasks that will serve as the example for the techniques presented here. This tutorial will cover the basics of defining an error function; choice of fit algorithm (e.g., particle swarm vs. gradient descent); evaluation of fit outcomes and choice of algorithm settings; a basic overview of distributed programming options in python (e.g., use of a parallel processing package vs. techniques based on “lock files”); and lastly, the use of the Python package “Cython” to port C code and libraries into python to decrease runtime for computationally expensive sections of code.

Corresponding author contact: rivkac@sas.upenn.edu

A computational cognitively-inspired technology for clinical diagnosis

Matthew Cook, University of Manitoba
Randall Jamieson, University of Manitoba

We used established computational models of semantic memory to classify peoples' mental health from free-form written reports. To make the point, we collected 500 undergraduates' written self-reports describing their mental health and scores on psychiatric instruments, converted the essays to a semantic hyperspace, and classified well and unwell by learning the structure of that hyperspace. Hyperspaces developed with several cognitively-inspired semantic models (i.e., LSA, BEAGLE, BEAGLE-RP) trained on meaningful text corpora (i.e., samples of text about mental health) classified novel essays better than chance. Consistent with work in medical diagnosis, the system performed best when classification of an unknown case reflected similarity to a small number of known instances. The goal of this work is to develop a fully automated classification system for accurate psychiatric diagnosis from natural language. The work contributes to an emerging body of evidence for natural language classification in computational psychiatry.

Corresponding author contact: cookm346@myumanitoba.ca

Integrating Language Modeling and Psycholinguistics

Melody Dye, University of California, Berkeley,
Michael Jones, Indiana University;
Marti Hearst, University of California, Berkeley

While there is now a substantial literature attesting to the importance of prediction in language processing, there has been surprisingly little interface between language modeling and psycholinguistics. Study materials are still frequently selected on the basis of either experimenter intuition, or on probabilities estimated from cloze tasks. When corpus data are used, the norm is to compute sequence probabilities on the basis of unsmoothed n-gram models; likewise, 'statistical learning' is typically limited to transitional probability estimation or simple co-occurrence tracking. Reliance on such a narrow range of statistics risks limiting the scope of inquiry to yet further proofs-of-concept. Information theory supplies a set of computational tools that may offer remedy. Depending on the object of investigation, measures of entropy and information can be computed over units of different granularity (such as syllables, morphemes, words, or larger sequences) and in relation to different properties (such as lexical, inflectional, or derivational). The source of the frequency distributions to which they are applied can also vary, ranging from word embeddings to probabilistic context free grammars. Indeed, these measures can be applied to the output of virtually any class of language model, including those that incorporate multiple converging sources of evidence (such as sequential weights or decay functions, or adaptations to the characteristics of the listener or channel). Comparisons across models of very different theoretical persuasions can be made, and individual model parameters adjusted to test key assumptions. This computational toolkit can thus be used to redescribe extant findings with greater precision and greater clarity, discriminate competing accounts, and suggest more nuanced empirical inroads.

Corresponding author contact: jonesmn@indiana.edu

Web/Experimental Tools

The Language History Questionnaire (LHQ) 3.0

Fan Zhang, Pennsylvania State University,
Anyu Yu, Pennsylvania State University,
Xiaowei Zhao, Emmanuel College
Ping Li, Pennsylvania State University

The language history questionnaire (LHQ) is an important tool for assessing the linguistic background of bilinguals or second language learners and for generating self-reported proficiency in multiple languages. Previously we developed a generic LHQ based on the most commonly asked questions in published studies (Li, Sepanski & Zhao, 2006) and provided researchers with a web-based interface (LHQ 2.0) that has flexibility in functionality, accuracy in data recording, and privacy for users and data (Li, Zhang, Tsai, & Puls, 2013). Since the publication of LHQ, hundreds of researchers have used it for their studies. These researchers have also provided many comments/requests/inquires for potentially new functions of the LHQ. The current version (LHQ 3.0) reflects our efforts to address some of the needs by the researchers. Among the many improvements we made to the web-based tool, the most important one is the addition of an aggregate score function, which can automatically calculate aggregated scores for a participants' second language proficiency level based on their inputs to several relevant LHQ questions. Such a score will help the researchers get a rough estimation/classification of different types of bilingual speakers. Furthermore, different researchers may be focused on different aspects and dimensions of a second language learner's language background. Therefore, LHQ 3.0 also provides researchers with the flexibility to assign different weights to different modules when calculating the aggregated score. The validity and reliability of this new function will be discussed.

Corresponding author contact: pul8@psu.edu

Presenting RDoC app – a web application for psychological assessment of patients

Yury Shevchenko, University of Mannheim
Felix Henninger, University of Koblenz-Landau

The research domain criteria (RDoC) is a framework for a new way of studying and classifying mental disorders. It focuses on measuring central psychological constructs that are related to systems of emotion, cognition, motivation and social behavior. The Central Institute of Mental Health (CIMH) in Mannheim has initiated the creation of an application that will assess the psychological state of patients of the CIMH hospital. At the SCIP conference, we are going to present the online application (RDoC app) that administers psychological assessment tools. There are five domains of the RDoC matrix that guide the development of the app: negative and positive valence, cognitive, social processes, arousal/regulatory systems. The current version of the app includes tests of working memory, attention, response control, information processing and reaction time. The RDoC app also contains administration tools to manage participants and tasks, analyze and visualize data. The web

application RDoC is programmed using JavaScript tools for both server-side and client-side environments: MongoDB, Express, Angular and Node. The psychological tests are built as reusable components with the lab.js experiment builder.

Corresponding author contact: yshevche@mail.uni-mannheim.de

Open-source and reproducible workshops: Creating portable learning environments to promote computational education in psychology and cognitive science

Alexandra Paxton, University of California, Berkeley

Aaron Culich, University of California, Berkeley

Jessica B. Hamrick, University of California, Berkeley

Christopher Hench, University of California, Berkeley

Todd M. Gureckis, New York University

Thomas L. Griffiths, University of California, Berkeley

New tools, methods, and analyses are being introduced to cognitive science and psychology at an incredible pace. While these advances open the door to greater productivity, more openness, and new research questions, a primary challenge facing the field surrounds questions of dissemination and education: Given limited time and financial resources, how can we improve access to cutting-edge advances to the broader research community? Being able to accommodate the full diversity of researchers -- spanning career stages, research areas, demographic characteristics, and computational skill levels -- is vital to the continued refinement of psychological theory and the wider adoption of the computational perspective. To that end, we describe here one step towards improving access to these opportunities: integrating tools of open science with open-source infrastructure software to create portable learning environments (PLEs) -- education ecosystems integrating tools like Jupyter Notebooks, JupyterHub, GitHub, Docker, Kubernetes, and Binder with cloud computing resources (like XSEDE Jetstream or Google Cloud). We report the successes and challenges that we faced during a recent workshop (Data on the Mind; <http://www.dataonthemind.org/2017-workshop>; funded by the Estes Fund) as an example of how researchers can use PLEs to create reproducible educational experiences that can provide lasting resources for our research community. Although our talk will focus primarily on informal educational opportunities, we will also discuss how these tools can be fruitfully applied to formal education.

Corresponding author contact: paxton.alexandra@gmail.com

Using Inquisit 5 for Online and Mobile Research

Sean C. Draine, Millisecond Software

Inquisit 5 is a powerful and easy to use platform for gathering psychological data in the lab, over the web, and on mobile devices. In this presentation, we'll demonstrate tools for creating custom, high-precision psychological tests that can be administered online and that adapt to the various form factors of modern connected devices. We'll also cover features like animation, drag & drop, drawing, and more. Come and see how Inquisit can "mobilize" your research.

Corresponding author contact: seandr@millisecond.com

Intelligent Tutoring Systems

Let's Walk About That: Random Walk Analyses Contextualize Performance in an Intelligent Tutoring System

Aaron D. Likens, Arizona State University
Kathryn S. McCarthy, Arizona State University
Laura K. Allen, Arizona State University
Danielle S. McNamara, Arizona State University

Learners generate an abundance of data while engaging with intelligent tutoring systems (ITSs). Conventional analytical approaches treat temporal variation in those data as random noise. However, dynamical systems theory and methodological approaches emphasize the importance of time in understanding behavior. Previous research has demonstrated that these dynamical approaches can be leveraged to uncover structure in ITS users' practice behaviors. This study builds on that research in two ways: (1) We used random walk analysis to reveal structure in ITS users' performance data. (2) We augmented random walk visualizations to better understand performance trajectories. Participants ($n = 149$) practiced with iSTART, a reading strategy ITS, across two sessions. Participants were randomly assigned to a game or nongame environment in session one and either stayed in this environment or switched in session 2. During practice, participants generated self-explanations (SEs) and received automated SE scores from 0-3. Augmented random walk trajectories, depicting SE scores over time, revealed distinct performance score patterns. To quantify these patterns, Shannon entropy was estimated for each participant's walk. Regression analyses indicated participants who switched environments had higher entropy than participants who did not switch. Results also revealed a quadratic relationship between entropy and average SE score such that entropy tended towards its minima for low and high performing participants. Together, the results stress that temporal variability is not random, but structured and meaningful. Further, these structures are sensitive to learning environment changes and predict overall performance. Discussion will focus on the promise dynamical approaches hold for understanding behavior.

Corresponding author contact: aaron.likens@asu.edu

Stairstepper: Adapting iSTART for Adult Literacy Learners

Kathryn S. McCarthy, Arizona State University
Cecile Perret, Arizona State University
Aaron D. Likens, Arizona State University
Amy M. Johnson, Arizona State University
Tricia A. Guerrero, University of Illinois at Chicago
Danielle S. McNamara, Arizona State University

There are few educational tools available for the nearly 2 out of 5 adults who read below functional literacy. To address this, we have begun adding features targeted at adult literacy learners to our existing intelligent tutor, Interactive Strategy Training for Active Reading and Thinking (iSTART). In the current study, we assess perceptions and performance of one of these features: StairStepper. While the main iSTART system uses complex scientific texts appropriate for high school students, StairStepper provides extended practice with texts and multiple-choice questions appropriate for adult literacy learners. Based on the learner's previous performance, StairStepper adapts in two ways: 1) amount of scaffolding (no self-explanation, prompting self-explanation, prompting self-explanation with feedback and revision) and 2) modifying the text difficulty level. These adaptive features are intended to support lower- and higher-order reading skills. In an initial pilot study of StairStepper, 52 undergraduates completed the Gates MacGinitie Reading Test (GMRT) and then received 45 minutes of iSTART training followed by 90 minutes of StairStepper practice. Overall, StairStepper training did not increase GMRT posttest score. However, further analysis of the in-system data showed that high-skilled readers quickly ascended in text difficulty while extremely low-skilled readers quickly descended. Readers in the middle (similar in GMRT score to adult literacy populations) showed greater variation in text difficulty. These data revealed a significant correlation between GMRT pretest score and rate of change in text difficulty during gameplay, $r(50) = 0.69$. Though preliminary, these results suggest that StairStepper is particularly appropriate for adult literacy readers.

Corresponding author contact: ksmccar1@asu.edu

Recognition of Structure and Meaning of Text in Struggling and Proficient Readers after Interacting with an Intelligent Tutoring System

Anne Lippert, University of Memphis
Brea Walker, University of Memphis
Raven Davis, University of Memphis
Qinyu Cheng, University of Memphis
Zhiqiang Cai, University of Memphis
Shi Feng, University of Memphis
Danielle Clewley, University of Memphis
Genghu Shi, University of Memphis
Arthur Graesser, University of Memphis

Comprehension is a critical component of literacy. Deep comprehension demands causal or goal-oriented reasoning and functional conceptual knowledge, while shallow comprehension entails recall of text features without necessitating a coherent understanding of the text. Compared to proficient readers, reading strategies of struggling adult readers are suboptimal when text requires deeper comprehension. To understand AutoTutor's effect on comprehension, we tested memory for text in students ($n=52$) and community members ($n=31$) after AutoTutor intervention. We first measured participant's literacy levels using the Test of Adult Basic Education (TABE) and then they completed two AutoTutor lessons. To assess recall for structure (shallow) and meaning (deep) of AutoTutor content, we used a four-alternative forced choice test. Specifically, participants were asked which sentence appeared in the AutoTutor lessons, and the correct choice was provided along with three distractors: text with the same meaning but different sentence structure than the actual text; text with the same

structure but different meaning than the actual text; text that differed both in sentence structure and meaning from the actual text. Deep comprehension is reflected in recall for the actual text or text with the meaning intact. We addressed the following questions: 1). Does AutoTutor promote comparable levels of deep comprehension across literacy levels? 2). Is comprehension affected by aspects of AutoTutor lesson text (e.g. difficulty)? 3). Do individual differences (e.g. gender) predict levels of deep comprehension after using AutoTutor?

Corresponding author contact: alippert@memphis.edu

Encouraging Deep Question Asking in Tutorial Dialogues Using a Wizard of Oz Technique

Colin L. Widmer, Miami University
Christopher R. Wolfe, Miami University
Christy Torrese, Miami University

Computer tutoring systems that interact with learners in natural languages dialogues are able to produce reliable and powerful learning gains. These systems can be designed to carry out many different forms of tutor-learner discourse. One form of discourse that has been successful is to encourage learners to generate questions, with the asking of deeper, more substantive questions that engage with the essential components of lesson material in particular leading to stronger learning. However, many learners find generating deep questions difficult and tend to generate shallower questions that focus surface details. To investigate the influence of tutor instructions and feedback on the types of questions learners generate and subsequent learning gains, a web-based Wizard of Oz system was used that permits researchers to simulate an intelligent tutoring system. After receiving a short lesson about breast cancer risk, learners were provided with instructions designed to encourage generating questions of various depths and then given the opportunity to ask questions, receiving answers through an animated avatar controlled by a researcher through the Wizard of Oz system. The tutor's instructions were found to have a large impact on the types of questions generated by learners, with learners instructed to ask questions that connected two or more pieces of lesson information together asking a greater number of deep questions than learners given other instructions. Additionally, the greater the number of deep questions a learner asked, the higher their performance was on multiple measures of learning outcomes completed after the tutor-learner discourse.

Corresponding author contact: widmercl@miamioh.edu

POSTER SESSION**(1) The time course of semantic and syntactic processing in Chinese word/sentence comprehension: Evidence from eye movements**

Xuran Zheng, Universidad Autónoma de Madrid
Lorena A. Martín-Arnal, Universidad Autónoma de Madrid
Ricardo Olmos, Universidad Autónoma de Madrid
José A. León, Universidad Autónoma de Madrid

In this experiment, we analyze the time course of semantic and syntactic processing when Chinese read sentences. Reader's eye movements were monitored following the paradigm used by Yang et al., (2009), where the relation between a single-character critical word and the sentence context was examined. Three conditions were developed: congruent, semantic violation, and semantic and syntactic violation. Thirty-five native Chinese speakers at Universidad Autónoma de Madrid took part in this study. They read all the 39 experimental Chinese sentences and their eye movements were registered in Eyelink 1000 plus. Two sets of analyses were carried out. First set was based on individual Chinese characters and including the critical word (CW), two characters before it (CW-2, CW-1) and three characters after it (CW+1, CW+2, CW+3). Second set was based on regions where pairs of two characters of interest were combined: pre-target region T-1, target region T, first post-target region T+1, second post-target region T+2. The eye movement data showed no differences in condition factor in the first-pass processing, only registered in CW respect to CW-2 and CW-1 in First fixation duration, as well as Gaze duration measures. However, there was a significant effect on condition in later processing, in regressions out in the semantic + syntactic violation comparing to congruent condition. These results suggest that the effects of a semantic + syntactic violation can be detected delayed by Chinese readers and that the processing of syntactic and semantic information is distinct in both first-pass and second-pass reading.

Corresponding author contact: ylzxr@hotmail.com

(2) Automated Essay Scoring (AES) with Latent Semantic Analysis (LSA) as an alternative to knowledge assessment in front of open-ended questions and multiple-choice tests

José Á. Martínez-Huertas, Universidad Autónoma de Madrid,
Olga Jastrzebska, Universidad Autónoma de Madrid,
Andrés Martínez, Universidad Autónoma de Madrid,
Lorena A. M. Arnal, Universidad Autónoma de Madrid,
Ricardo Olmos, Universidad Autónoma de Madrid,
José A. León, Universidad Autónoma de Madrid,

Multiple-choice testing has many advantages (e.g., low-cost or content validity), while it also has some limitations that can be avoided by using open-ended questions (e.g., different cognitive demands or deeper understanding) (Shapiro, & McNamara, 2000). This study presents LSA's Inbuilt Rubric (IR) as a

new approach to knowledge assessment. IR method transforms the latent semantic space into a new meaningful one where a human expert rubric is mapped in this new space (Olmos et al.,2016). This method can detect subtle student contents in their essays in front of classical LSA's assessment methods like Golden Summary (GS). METHOD: In this study, 100 undergraduate students read two texts. For assessing the quantity and quality of acquired knowledge, students summarized the texts and answered two multiple-choice tests. Summaries were scored by three independent human experts from the previously established rubrics. RESULTS: In text 1, Pearson's r between expert rubrics and multiple-choice tests were .74. IR correlated .81 with rubrics and .75 with tests, while GS correlated .61 and .59, respectively. In text 2, rubrics and multiple-choice correlated .70. IR correlated .78 with rubrics and .49 with tests, while GS correlated .60 and .34, respectively. Likelihood ratio test showed that IR and GS Pearson's r were statistically different in all texts and methods. CONCLUSION: AES seems to be a good alternative to knowledge assessment in academic contexts because of its favorable characteristics. Particularly, IR is a relevant alternative to open-ended question's rubrics and multiple-choice tests because this method maintains the best of both assessments.

Corresponding author contact: joseangel.martinezhuelas@gmail.com

(3) Mapping the Theories of Cognitive Science Using LSA

Pablo Andrés Contreras Kallens, University of California, Merced
Rick Dale, University of California, Los Angeles

The theoretical structure of Cognitive Science has shifted dramatically from what it was in the heyday of traditional representational and information-processing accounts of the mind. The appearance of alternative approaches, usually categorized under the term “Embodied Cognition”, has altered the perceived consensus of the discipline and cast doubt upon what were once its fundamental assumptions. Prior meta-theoretical studies of Cognitive Science have traditionally focused either on qualitative research based on extensive literature reviews, or on quantitative explorations of non-theoretical features like the balance of disciplines in core journals, or the topics to which the different theories refer. In contrast, in this study we tested the effectiveness of a quantitative tool –Latent Semantic Analysis– in capturing the boundaries between different theories using the data provided by word use in abstracts obtained from Web of Science. The results of our analysis suggest that LSA can be used to represent these different theories. First, theories have relevantly different features. Secondly, the similarity in word use of the abstracts of the different theories mirrors the intuitive narrative of the theoretical ridges in Cognitive Science. Finally, we show that word use can be a reliable predictor of theoretical inclination when using a statistical representation of the patterns of word use of the abstracts of the different theories.

Corresponding author contact: pcontreras3@ucmerced.edu

(4) Complex Communication Dynamics: Exploring the Structure of an Academic Talk

Camila Alviar, University of California, Merced
Rick Dale, University of California, Los Angeles
Alexia Galati, University of California, Merced

Communication is a multimodal phenomenon. There is ample evidence that words, gestures, pitch variations, external objects, all help shape the meaning of what is being communicated. However, little is known about the cognitive mechanisms supporting multimodal coordination, especially during natural language use. We explored a natural dataset of real academic lectures to determine how different communication modalities are individually used and dynamically combined during the presentation of complex information. We used both automated and semi-automated techniques, such as frame differencing analysis and automated speech processing, to extract behavioral measurements from video. The dynamics of body movement, slide change rate, speech rate, articulation rate, fundamental frequency, and intensity of 30 speakers were analyzed. The results of the mixed-effects models using time as a linear and a quadratic predictor reveal some consistent but statistically subtle patterns in the use of speech rate, articulation rate, intensity, and body motion across the presentation. Using principal components analysis, we observe patterns of system-like covariation between modalities, though the observed patterns are more difficult to interpret. These results do suggest that the cognitive system is integrating body, slides, and speech to act in a coordinated manner during natural language use. Further research is needed to clarify the specific coordination patterns that occur between the different modalities.

Corresponding author contact: malviarguzman@ucmerced.edu

(5) Evaluating methods of keyword extraction from short comprehension passages using word embeddings

Prudhvi Raj Dachapally, Indiana University
Fateme Torabi Asr, Indiana University
William, R. Aue, Purdue University
Jeffrey D. Karpicke, Purdue University
Michael N. Jones, Indiana

We present a novel approach to unsupervised keyword extraction from short passages. This method is integrated with a tutoring system employed in elementary classrooms for educational research, in particular, to study how information presentation (e.g., highlighting keywords in text) improves learning. Previous methods for automated keyword extraction work at the corpus level or on large documents, where word co-occurrence can directly be leveraged for identification of concept graphs or importance of a word in one document relative to others. However, these methods do not easily scale to a situation where a single short passage is given. We propose a novel approach based on the assumption that keywords in a short document should be closely related in meaning. We use word2vec embeddings trained on gigabytes of raw text to cluster words in a document and then select the cluster with the best set of keywords. Two clustering methods are employed, one based on distances between all candidate words to cluster centers and the other based on integration of previous predictions and the centroid. Three strategies are proposed for best cluster selection, including one that incorporates TF-IDF scores. Our clustering approaches outperform previous methods such as TextRank and classic TF-IDF in extracting keywords from short passages, where elementary school students who read the passages to answer comprehension questions define the gold standard. Our system is also very efficient for real-time application given the fact that both word embeddings and TF-IDF scores can be pre-trained on an arbitrary large corpus such as Wikipedia.

Corresponding author contact: torabiasr@gmail.com

(6) Fatherese: Utilizing Corpora to Analyze Male Caregiver Driven Childhood Language Acquisition

Dawn Carroll, California State University, Fullerton

Through the linguistics sub-field of Child Language, researchers performed several studies on mother-child interactions. The overall objective of these studies was to better understand the way that mothers speak to their children and how their children can acquire their L1, or first language. Despite the immense amount of knowledge acquired from these previous studies, researchers have not thoroughly investigated male caregivers' speech, or fatherese, and how it can affect the way that a child uses their language. As a result of this limited research, I chose to analyze features of fatherese in comparison to motherese and how it affects childhood L1 acquisition. I reviewed the conversations of four children under the age of five and their separate interactions with both their mother and their father from the Gleason Corpus within the CHILDES (Child Language Data Exchange System) Corpus. From analyzing the corpora, I discovered a male driven influence to how children interact in respect to their perceived gender norms.

Corresponding author contact: dcarr010@csu.fullerton.edu

(7) EFashiny: A shiny application for exploratory factor analysis

Chi-Lin Yu, National Taiwan University,
Ching-Fan Sheu, National Cheng Kung University

We develop a shiny application, EFashiny, for exploratory factor analysis (EFA). The graphical user interface (GUI) is designed to free users from scripting in R by wrapping together various packages for data management, factor analysis, and graphics. EFashiny provides visualization of factor loadings with confidence intervals to facilitate proper interpretation of extracted factors. Results of analysis in tables and graphs are presented on-line and can be exported. Features of EFashiny are demonstrated with two real data examples.

Corresponding author contact: csheu@mail.ncku.edu.tw

(8) Median Split and Type 1 Error Inflation

Bradley Smith, University of Manitoba
John R. Vokey, University of Lethbridge
Randall K. Jamieson, University of Manitoba

Median split is a common method used to draw conclusions in psychology. In that procedure, participants are scored on two variables, m and x (e.g., IQ and short term memory span). Then, groups are formed based on scores on m (e.g., IQ) to conduct an analysis of differences in scores on x (i.e., short term memory span). Although median split techniques appear frequently in published work, that strategy encourages severe type 1 error inflation (i.e., reporting a difference between groups where no difference exists). The issue is important in light of recent debate over the reliability of conclusions drawn in psychological research. We address the issue by presenting Monte Carlo simulations that show the severity of type 1 error inflation conditional on a median split, we provide a R package for

analysis and correction of error inflation, and we discuss the difference and proper use of abductive versus deductive inference in scientific and experimental reasoning.

Corresponding author contact: smithb21@myumanitoba.ca

(9) Simulating learning in a T-maze with a model of memory: Embodying a psychologically-inspired theory of learning

Bradley Smith, University of Manitoba
John Vokey, University of Lethbridge
Randy Jamieson, University of Manitoba

We implemented MINERVA-AL (Jamieson, Crump, & Hannah, 2012), an exemplar based theory of learning, in a Pololu m3pi robot. We tested the robot's learning in a simple T-maze task. Unlike many machine learning algorithms, the exemplar method assumes that each learning trial is recorded as a trace in memory and that behaviour follows from cue-driven similarity-based parallel retrieval from the database. The approach circumvents traditional problems including exploration vs. exploitation, maximizing reward signals, a traditional value function, and other complications identified by Sutton and Barto (1998). The demonstration demonstrates the claim that learning is a corollary of exemplar-based storage and retrieval from memory and shows how an exemplar theory of learning can be coded, implemented, and executed in a behaving robot.

Corresponding author contact: smithb21@myumanitoba.ca

(10) Applying convolutional neural networks to fMRI data classification: A deep learning approach

Jinlong Hu, South China University of Technology, Guangzhou, China.
Bin Liao, South China Agricultural University, Guangzhou, China.
Xiaowei Zhao, Emmanuel College, Boston, MA, USA.
Ping Li, Pennsylvania State University, University Park, PA, USA.

Convolutional neural network (CNN) is a type of deep learning artificial neural network with convolutional and feed-forward operations of two or more hidden layers. CNN has been successfully applied to visual recognition tasks (see Kriegeskorte, 2015 for a review). In this presentation, we apply CNN to functional Magnetic Resonance Imaging (fMRI) data to classify a large number of time-series brain imaging data obtained from a naturalistic text reading task in the MRI scanner. In order to increase accuracy, several specific generalization strategies and training optimization methods are examined for the classification, along with some practical considerations for implementing and training the CNN models. We experimentally compare our CNN approach with several existing deep neural network methods to verify the effectiveness of the proposed CNN approach. This represents an initial effort in using CNN to analyze fMRI data in a cognitive task involved in reading comprehension.

Corresponding author contact: jihu@scut.edu.cn

(11) Speech Emotion detection based on Hilbert-Huang transform (HHT) and deep learning

Zhi-Yong Tao, Harbin Engineering University

Ya-Xian Fan, Harbin Engineering University

Bo-Jie Zhou, Harbin Engineering University

Xiaowei Zhao, Emmanuel College, Boston

Speech emotion detection has been an attractive subject in various fields such as human-computer interactions, psychological monitoring, psychotherapy, and AI. However, it is still a challenge to effectively distinguish various emotions from vocal signals of human speech. Especially, what part of a vocal signal may best represent emotions? To answer this question, we implemented the Hilbert-Huang transform (HHT), a novel signal processing technique, to the Mandarin Chinese vocal emotional stimuli database (Liu & Pell, 2012). Different from traditional Fourier transform based analyses, HHT is designed to analyze nonlinear and non-stationary signals, which are the natural situation of human speech. Following HHT, we first obtained fourteen Intrinsic Mode Functions (IMFs) from the original signals based on the so-called Ensemble Empirical Mode Decomposition (EEMD) algorithm. Then, Hilbert Transform was applied to each IMF component to extract their inherent information, such as their changing instantaneous frequencies. Finally, the extracted information was input into two unsupervised learning models: a Self-Organizing Map (SOM, Kohonen, 2000) and a Deep Belief Network (DBN, Hinton, 2006) with three hidden layers. The purpose was to investigate which IMF(s) can provide reliable information to classify original vocal signals into correct emotion categories. Our results based on SOM indicated that the emotion information was not evenly distributed to every IMF and our second IMF provided better recognition rate of emotions than other IMFs. To understand why it was the case, we further analyzed the IMFs' internal representations on the three hidden layers of the deep-learning based DBN.

Corresponding author contact: zytao@hrbeu.edu.cn

(12) Detecting Contradiction in Agent Source Monitoring During Expository Text Comprehension

Shi Feng, University of Memphis

Zhiqiang Cai, University of Memphis

Arthur C. Graesser, University of Memphis

Contradiction detection has been shown to be important in reading comprehension. Recent studies have found that source monitoring is an important reading strategy to resolve contradiction between texts. The current study investigates the effect on comprehension in an Intelligent Tutoring System when agent contradiction and agent source monitoring are incorporated into the learning environment. Participants read three texts in a page-by-page format and listened to two agents speaking statements that either contradicted the text or were consistent with the text. The contradiction manipulation was in four conditions: teacher contradiction, student contradiction, both contradiction and no contradiction. A source monitoring recognition exam was conducted after participants read each text. Participants' comprehension was also evaluated by an inferential comprehension exam that was tailored to each page. Consistent with previous findings, participants are biased towards responding the teacher agent as the source, even though participants remember a

contradiction best when the contradiction is repeated by each agent. However our results suggest that agents relaying contradictions alone does not affect comprehension. The results are explained by the theories of social contagion and models of comprehension.

Corresponding author contact: sfengmemphis@gmail.com

(13) The Semantic Cognition in Language (SCiL) Application: Expanding Cognition's Research Frontiers with iOS

Adrian Cunningham, University of California, Riverside
Felipe Ranzani de Luca, Centro Universitário Senac - Santo Amaro, São Paulo - Brazil,
Curt Burgess, University of California, Riverside
Bruno Moscardi Flora, Centro Universitário Senac - Santo Amaro, São Paulo - Brazil,
Kimberly Miller, University of California, Riverside
Ashima Kumar, University of California, Riverside
Melissa Saidak, University of California, Riverside
Meera Shah, University of California, Riverside
Giselle Urquijo, University of California, Riverside
Mabel Wong, University of California, Riverside

The Semantic Cognition in Language (SCiL) project is a collaboration between the Computational Cognition Lab, University of California, Riverside, Apple Inc., and an Apple-sponsored developer team at the Centro Universitário Senac - São Paulo, Brazil. The goal of this project is to develop an application for iOS and Mac devices that can be used to conduct cognitive experiments to investigate normal and neurologically impaired memory and language function. SCiL features millisecond measurement precision, high customizability, and a browser-based application for creating experiments. Participants will be able to download the SCiL app from the App Store and browse available experiments. The implemented cognitive procedures include the lexical decision priming task and the categorization task. The app also has a module for the presentation of questionnaires. A questionnaire or similar assessment tool can be upload by the experimenter. The SCiL project has established an Assessment Repository that includes many demographic, clinical and social psychological tests. SCiL will be HIPAA compliant to provide reliable security for storing and retrieving sensitive user data. To establish the reliability of the app, we will present the results of an experiment designed to replicate the Lund, Burgess and Atchley (1995) semantic priming experiments with three types of word relationships and extend this research by evaluating the role of vocabulary level in the priming effect. The long term plan is to develop a community of researchers who will use the site for collaborative projects that will capitalize on this technology.

Corresponding author contact: curt@ucr.edu

(14) iPad applications for assessing and training attention function

Toshio Ohyanagi, Sapporo Medical University
Kunihiro Kanaya, Sugata Hospital
Yasuhito Sengoku, Sapporo Medical University
Lili Liu, University of Alberta
Eleni Stroulia, University of Alberta
Masako Miyazaki, University of Alberta

Using tablets and smart phones is becoming popular in both research and clinical settings. We have developed an application to run three reaction time tasks and a Bluetooth Low Energy device running with an iPad to assess attention function of older adults and people with inattention (Ohyanagi et al., 2015). Recently, we developed two new game applications to train attention function of the subjects. One is so-called a “Whack-A-Mole” game originally developed as a mobile (Android) game to help assess decline in cognitive function (Tong, Guana et al., 2015). We modified the original one to adjust the degree of difficulties by adding distractors and Go/no-Go feature. The other one is EyeExercise which to train eye movements of the subjects by using JiNS MEME. The JiNS MEME is an eyeglass type equipment to detect eye and head movements by using electrooculogram (EOG) signals and motion sensors. This device is quite unique and its clinical use has to be further explored for older population. In this presentation, we will describe the details of our new games and preliminary results of using them for 12 healthy older adults and 8 patients. As a result, five patients could not play either Whack-A-Mole or EyeExercise and other patients could play both or one of the games but the scores were lower than those of healthy older adults. On the other hand, all healthy older adults could play Whack-A-Mole well and eleven of them also could play EyeExercise well.

Corresponding author contact: ohyanagi@sapmed.ac.jp

(15) Lexical Association and False Memory for Words: A Reexamination Ten Years Later

Yuh-shiow Lee, National Chung-Cheng University

Languages change over time. The enormous database from the internet provides a valuable tool to investigate this change. This study examined the relationship between lexical association and false memory produced by the DRM paradigm and compared this relationship between data collected from two different time points. The co-occurrence rate, as an index of lexical association, of two Chinese character was calculated based on data from the Internet. The result was compared with that collected 10 years ago based on a collection of 751 million paragraphs from more than 10 million Chinese Internet web pages and correlated with false recall and recognition data of Mandarin-speaking participants collected more than a decade ago. The original co-occurrence rate was significantly correlated with false recall and recognition of Mandarin-speaking participants, and not of English-speaking participants (Lee, Chiang & Hung, 2008). Based on the more recent collected co-occurrence rate, the correlation for Taiwanese participants remained significant. Moreover, there was also a significant correlation between the two co-occurrence rates collected 10 years apart. Results are discussed in terms of the stable relationship between lexical association and false memory for word lists. Reasons for this stability and limits of the study are also discussed.

Corresponding author contact: psyysl@ccu.edu.tw

(16) Does question format change inferences about corrected false information? Comparing elicitation formats in the continued influence effect paradigm

Saoirse Connor Desai, City, University of London

Stian Reimers, City, University of London

The reliability of the continued influence effect (CIE) paradigm was examined using a novel methodology. CIE studies consistently show that information presumed true at encoding but later turns out to be false continues to be influence memory and reasoning. In the standard task, participants read an event report containing false information and a subsequent correction. Inferential reasoning about the event is then measured through a series of open-ended questions. A correction is considered effective to the extent it reduces reference to false information compared to no correction. Theoretical accounts explain the CIE either as a faulty mental model of the reported event or a failure to retrieve the context false information. Previous research has not addressed whether the CIE reflects demand effects related to ambiguity in open-ended questions. To address this we compared responses to open- and closed-ended questionnaires under different types of correction in two web-based studies. Results showed a continued influence effect of false information in response to both open- and closed-ended questions. Reference to false information in response to inference questions was equally likely whether the information was corrected, or not. The findings have important methodological implications. First, the CIE is observed using different elicitation formats. Second, it is feasible to run memory-based experiments on the web and elicit responses through open-ended questions (i.e., participants give reasonable answers). Third, the continued influence effect generalizes to more heterogeneous populations than university subject pools. These findings strengthen the reliability of results previously obtained using the continued influence effect paradigm.

Corresponding author contact: saoirse.connor-desai@city.ac.uk

(17) Do Constructive Play and Exploratory Representational Play in Video Games Effect Problem Solving and Divergent Thinking?

Elizabeth M. Cedillos-Whynott. Miami University

Video game play has been shown to improve such cognitive processes as reaction time, attention, and perception. However, there are few studies examining the effect video games have on higher order cognitive processes, specifically how play itself is involved. The goal of this study is to examine whether constructive play (CP), crafting and building tools and structures, and exploratory representational play (ERP), exploration of the environment, can be shaped in the video game "Don't Starve." Furthermore, I aim to determine whether CP and ERP influences problem solving and divergent thinking, respectively. College-aged participants will be randomly assigned to one of three conditions: CP, ERP, or control. The CP condition will have an increase in the amount of resources available to encourage crafting and building structures, along with objectives which encourage crafting and building. The ERP condition will have a decrease in the amount of resources available, requiring participants to explore their environment, along with objectives which encourage collecting resources. Control participants will read a scientific article and answer 5 questions about the article. After one hour, participants will complete two problem solving measures: The Tower of Hanoi puzzle and the River Crossing task, and two divergent thinking measures: The Alternate Uses task, and the Instances Assessment of Creativity.

Data will be presented as to whether designed video game play, CP and ERP, can influence problem solving and divergent thinking. Differences between CP, ERP, and control, on the four assessment measures will be discussed.

Corresponding author contact: cedillem@miamioh.edu

(18) Walk and Talk: The impact of attentional load on gait dynamics in young adults

Simon Ho, University of British Columbia
Kash Daud, University of British Columbia
Ute Leonards, University of Bristol
Todd C. Handy, University of British Columbia

Many studies have shown a relationship between executive function and gait in elderly populations, however, less is known about the nature of that relationship in young adult samples. University-aged participants were asked to complete the verbal trail making test while walking down a hallway. Gait dynamics were recorded using smartphone-based accelerometry. Findings suggest that gait is negatively impacted when attentional resource competition is high, even in young adult populations.

Corresponding author contact: simon@psych.ubc.ca

(19) Online Survey Tools in Longitudinal Cognitive Psychology Research: A Methodological Review

G. Kyle Gooderham, University of British Columbia
Simon Ho, University of British Columbia
Todd C. Handy, University of British Columbia

The intersection of health and cognitive psychology provides a fertile ground for the application of computer technologies. In our research we investigate the relationship between exercise and cognitive performance, with specific focus on the effects of intensity, duration, and type of exercise on various cognitive functions. The current study examines seven-day exercise behaviours, correlating self-reported daily physical activity with cognitive function and retrospective recall of physical activity. A significant limitation to researchers is the trade-off between accuracy and availability of current models of physical activity trackers, and therefore, alternative methodology must be used. In this methodological review we examine the challenges of collecting and reconciling self-reported activity data across longitudinal and cross-sectional study designs without the aid of physical activity trackers, and how online survey tools can be leveraged to mitigate these issues. Difficulties that must be addressed in all longitudinal designs are viewed in the context of online data collection, with solutions offered that proved fruitful in the current study. Additionally, we review how quantitative self-reported physical activity data are effectively measured using electronic tools and what additional instruments may be utilized to strengthen the findings of this and further efforts in this research domain. Finally, we offer researchers and programmers the requirements for a program that will be necessary to diminish these challenges in future research.

Corresponding author contact: kyle@psych.ubc.ca

(20) Modeling the moral landscape of political affiliation in the United States

Karie Moorman, UC Merced

Jeff Yoshimi, UC Merced

Political affiliation is a significant and effective predictor of people's attitudes and reasoning about many social problems (e.g., climate change, immigration, poverty). How our embodied experiences define our personalities and moral traits, they too define our political identities, and this influences the ways in which we define social groups, how social groups interact, and how we define and re-define social policies in the United States (e.g., Cunningham-Parmeter, 2010; Lakoff, 2002; Santa Ana, 2002). This suggests that our political attitudes are associated with specific cognitive mechanisms and tendencies to categorize and contrast social groups and social problems in certain ways (e.g., morally, Wehling, 2013; as opposing categories, Heit & Nicholson, 2010). Thus, when we reason about the political attitudes of others, we tend to draw on the same moral categories we use to define ourselves (e.g., Lakoff, 2008; Lakoff & Wehling, 2016). The purpose of this paper is to model this phenomenon using data collected from the Moral Political Family Scale (MPFS) survey (Wehling, 2013) and neural networks. One theory of how humans represent categories is via neural representations that arise via incremental tuning of synaptic connections using some form of unsupervised learning (Rumelhart & Zipser, 1985; Kohonen, 1990). These mechanisms produce representational spaces that preserve the similarity structure of an input space (e.g., banks of abstract neurons or nodes in which nearby nodes represent similar stimuli). We model the MPFS data using a self-organizing map in Simbrain (Tosi & Yoshimi, 2017), a tool designed in part to facilitate improved pedagogy and research in the areas of neural networks and cognitive science. The map can be viewed as a hypothesis about the way someone exposed to this data might represent the political affiliations of others.

Corresponding author contact: kmoorman@ucmerced.edu

PRESIDENT'S SYMPOSIUM**Predicting Human Learning with Recurrent Neural Networks**Michael C. Mozer (mozer@colorado.edu)Institute of Cognitive Science and Department of Computer Science
University of Colorado, Boulder

Psychological models of learning and memory have traditionally been used to explain experimental outcomes and to guide experimental design. In principle, models can also be used to predict an individual's future performance following a specific sequence of training trials. Models that accurately predict performance can be embedded in surrogate-based optimization procedures to determine training trial sequences that ought to maximize learning gains. We explore the use of recurrent neural networks (RNNs) for performance prediction in the domains of cognitive skill acquisition and category learning. The hidden layers of an RNN represent the learner's evolving knowledge state. We find that these generic, flexible models often do not significantly outperform psychologically grounded models. However, we argue that their true value will be revealed as we collect larger data sets and richer encodings of the stimuli and additional features describing learners and their environment. (This work was performed in collaboration with Mohammad Khajah, Karl Ridgeway, and Brett Roads.)

Neural Dynamics of Spontaneous Thought: Evidence from fMRI and EEGKalina Christoff (kchristoff@psych.ubc.ca)Department of Psychology
University of British Columbia, Vancouver

Mind-wandering has recently come to occupy a central position in cognitive psychology and neuroscience. Most theories and research so far have examined it in terms task-unrelated or stimulus-independent mental contents that occur at particular moments of time. A defining feature of mind-wandering, however, are its dynamics: how thought moves over time. In this talk, I will introduce a dynamic framework for understanding mind-wandering and its neural basis. I propose that mind-wandering is best understood as a member of a larger family of spontaneous thought processes – a family that also includes creative thought and dreaming. I will also discuss recent studies from my lab using fMRI and EEG to investigate the fine-grained temporal dynamics of brain activity underlying spontaneous thoughts. In these studies, we employed the first-person reports of experienced meditators to index the onset of spontaneous thoughts, and examined brain activity preceding indications of spontaneous thought onset. Across two studies, we found a consistent temporal trajectory of large-scale brain networks leading up to spontaneous thought reports that may map onto distinct mental processes such as thought generation, crystallization, and appraisal.

Deep Learning of Science Concepts from fMRI data of Naturalistic Reading

Xiaowei Zhao (zhaox@emmanuel.edu)

Department of Psychology

Emmanuel College, Boston

Neural networks based on deep learning algorithms have gained popularity in the last few years in computer vision and natural language processing. However, it is not until recently that researchers have begun to apply Deep Neural Networks (DNN) on large-scale neuroimaging data, and the DNN applications in the neurobiology of language remain scarce. In the current study, we applied a multi-layer deep neural network to an fMRI dataset collected from 50 native English speakers who performed a naturalistic reading task in which five short science texts were presented sentence by sentence. Our fMRI volumes were collected based on a novel technique that allows the collection of both BOLD signals and eye-movement data and the analysis of fMRI data using the onset of first fixation on a word as time window for modeling the hemodynamic response for specific words during naturalistic reading (Schuster et al., 2016; Desai et al., 2016). The fMRI image data were preprocessed, segmented into anatomical regions of interests based on the AAL brain atlas, and used as input in a DNN with three layers. The network was trained to complete different classification tasks including different topics of the text (math, engineering, space, environment, or technology) and different type of words (concrete vs abstract).

(This work was performed in collaborated with Chun-Ting Hsu, Benjamin Schloss, JinLong Hu and Ping Li.)

KEYNOTE**Understanding the Language-Learning Brain with Cyber-Enabled and Computational Methods**

Ping Li (Pul8@psu.edu)

Department of Psychology and Institute for CyberScience
Pennsylvania State University

How does the learning of a new language in adulthood lead to brain changes when the brain is already committed to one's native language? A major research question in the past has been how specific linguistic experience shapes brain function and structure, but little work has focused on cyber-enabled language learning and the impact it brings upon the mind and the brain. Using data-driven analytics we examine not only how second language experience induces functional and neuroanatomical changes, but also whether these changes can occur as a function of learning contexts and how cyber-enabled methods can bring about positive changes in different learners. Our research integrates computational modeling and data-intensive brain science, and addresses important questions about neuroplasticity, individual differences, and knowledge representation.

Corresponding author contact: pul8@psu.edu

.....Notes.....

Notifications

Birnbaum Scholarships

If you are a first-time graduate student attending SCiP, and the primary presenter, you may be eligible to have your registration waived as part of the Birnbaum Scholarship for first-time graduate presenters. Please contact our secretary, [Rick Dale](mailto:rdale@ucla.edu) (rdale@ucla.edu) for details.

Castellan Award

The Society sponsors The Castellan Student Paper Award for the outstanding student paper or poster. Student papers or posters on the application of computers to any area of psychology (theoretical, experimental, and applied) are welcome. Eligibility is open to work done by a student currently enrolled in undergraduate or graduate courses, or work done as part of a course, thesis, or other student research by a person who graduated within the last year. The student must be the primary author and the presenter of the paper.

The winning author will receive a one-year membership in the Society, a one-year subscription to BRM, and a \$200 cash prize. Students who wish to have their papers considered for the award should indicate their desire at the time of submission and are required to submit the complete manuscript for evaluation for the competition by October 15, 2017. The award for the outstanding student paper will be presented at the conference during the business meeting.

The 2016 winner of the Castellan Award was Pascal Kieslich:

Pascal Kieslich, University of Mannheim, Germany

Winning Paper: "Mousetrap: An integrated, open-source mouse-tracking package."

Special Issue of Behavior Research Methods

Conference presenters (both spoken and poster presentations) are invited to submit articles based on presentations at the SCiP meeting to a special issue of Behavior Research Methods, the official journal of the SCiP. The September 2018 issue of Behavior Research Methods will feature the 2017 SCiP conference. Although papers submitted to the special issue go through a normal review process, the time to publication is greatly accelerated. Consequently, this is an excellent opportunity to get your important research published in an extremely timely fashion. This is one of the important perks of the attending the conference. Submissions should be prepared according to the guidelines of BRM, which can be found at <http://brm.psychonomicjournals.org/>, and submitted through the BRM submission portal. It is important to indicate that the submission is for the "SCiP Issue."

The deadline for article submission is December 31, 2017.



**PSYCHOLOGY
SOFTWARE TOOLS**

Celebrating 30 Years

-Prime 3

**Use E-Prime with Chronos
for Millisecond Accuracy**



**9 - 22ms
latency**



< 1ms



**10 - 19ms
latency**

CHRONOS

A Multifunctional Response and Stimulus Device

pstnet.com

