

SCiP 2012 Program



Minneapolis, Minnesota

Thursday, November 15, 2012

Welcome Message from the SCiP President

Welcome to the 42nd annual meeting for the Society for Computers in Psychology (SCiP)! I hope you will agree that we have an extraordinary program for you this year.

Over the last four decades SCiP has played an important role in psychological science by promoting the use of computers in both research and education. SCiP's role may be further enhanced today given (1) the widespread use of the internet for learning and innovation and (2) the availability of web-based technologies and other cyber tools for teaching and research. This new role of SCiP is clearly reflected in this year's conference, in the Keynote Speech by Brian MacWhinney, the Presidential Symposium, and the Symposium on New Developments in Web-Based Data Collection, among many talks and posters.



Given the internet era, the challenge to members of our society is clear: how can we take advantage of the enormous potential that the web offers us for promoting psychological science and education, and how can we make the best use of our collective talents to push the SCiP mission to the next level? As a minor step and an example for answering such questions, I have worked in the past year with other SCiP colleagues to revamp our website, so that <http://scip.ws/> takes a new look, having user-friendly functionality and easy-to-navigate menus with relevant information. Your comments and suggestions on how to enhance our society's website are certainly welcome.

Highlighting the role of the web does not suggest we ignore other fine traditions that SCiP has set in the past. We continue to see that the SCiP community provides creative solutions to learning, teaching, and research through innovative technologies including software and hardware development, analytic tools for large-scale corpora or datasets, tools for optimizing experimental design and data analyses, and computational and semantic space models of human cognition and behavior. All these topics are well represented in this year's conference, in both the spoken papers and the posters. Following the tradition set by Joe Magliano last year, I have decided to hold a Presidential Symposium on Innovative Technologies and the Science of Learning, featuring speakers who use a variety of methods and perspectives to study natural language using intelligent tutors, eye-tracking technology, and data mining techniques and cognitive models (Hu, Perfetti & Koedinger, and Yu).

I look forward to seeing many old and new friends at the SCiP conference. If you are new to the society, please do not hesitate to ask me or members of the SCiP Steering Committee questions or to provide your views and opinions of the society. Finally, all presenters, whether of spoken papers or posters, are invited to submit manuscripts based on your presentations to the Special Issue of *Behavior Research Methods* (see <http://scip.ws/> for details).

Enjoy SCiP 2012 and your visit to Minneapolis!

A handwritten signature in black ink, appearing to read 'Ping Li'.

Ping Li

Table of Contents

Welcome Message.....	2
General Information.....	3
About SCiP	
SCiP Officers	
Presentation Guidelines	
Conference Location	
Conference Maps	5
Conference Schedule.....	6
Poster Presentation List.....	7
Keynote Speech.....	8
Presidential Symposium.....	9
Session A	10
Session B	12
Session C	15
Session D.....	17
Session H.....	19
Session I.....	21
Poster Session (Session F)...	23
Notes (for yourself).....	35
Exhibitors Information.....	37

General Information

About SCiP

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 set a special goal of aiding psychologists in using microcomputers in their teaching and research. We have also encouraged consideration of the psychological aspects of hardware and software development and design. Membership is open to any person who has an academic degree and who is active in scientific applications of computers to psychological research.

SCiP Officers

SCiP 2012 is organized by the SCiP officers and steering committee members listed below. Please contact them if you have any problems during the conference.

- President: Ping Li
Pennsylvania State University
pul8@psu.edu
- President Elect: Michael Jones
Indiana University
jonesmn@indiana.edu
- Past President: Joseph P. Magliano
Northern Illinois University
jmagliano@niu.edu
- Secretary-Treasurer: Rick Dale
UC Merced
rdale@ucmerced.edu

Steering Committee

- Patrick Conley,
University of Wisconsin
pconley@uwsp.edu
- Rick Dale,
UC Merced
rdale@ucmerced.edu

- Randy Jamieson
University of Manitoba
jamiesor@cc.umanitoba.ca
- Chris Koch
George Fox University
ckoch@georgefox.edu
- Kay Livesay
Linfield College
klivesa@linfield.edu
- Otto MacLin
University of Northern Iowa
otto.maclin@uni.edu
- Danielle McNamara
Arizona State University
dsmcnamara1@gmail.com
- Douglas Mewhort
Queen's University
mewhortd@queensu.ca
- Ulf-Dietrich Reips
University of Deusto
u.reips@ikerbasque.org
- Yana Weinstein
Washington U St. Louis
y.weinstein@wustl.edu
- Xiaowei Zhao
Emmanuel College
zhaox@emmanuel.edu

Conference Program

The conference program contains the conference schedule, the abstracts of all presentations and a list of authors. Please copy this program to your laptop or iPad/iPhone. Hardcopies of this program will be made available to all registered participants of the conference.

SCiP Time

To allow people to visit presentations in different sessions, we need to keep a tight schedule. We will time each presentation strictly with signs (5 min, 1 min, stop). Please observe the signs.

Presentation Guidelines

Talks

For oral papers, presentation time will be limited to a total of 15 minutes, which includes discussion time. The conference site provides a laptop (PC). If you want to use your own laptop, make sure it has a VGA connector (and bring your own connector or adapter!). Copy your presentation file to a jump driver just in case your own computer does not work with the setup.

Posters

The poster session will be held from 1:00pm to 2:30pm in Ballroom A of the Convention Center. Each poster will be assigned a presentation number, and please mount your poster to the poster board according to assigned number between 12-1pm. See Page 7 for details.

Keynote Speech

Dr. Brian MacWhinney
Carnegie Mellon University

“Using the Web to study real-life language learning: measures, experiments, models, tutors, and tours”

Location

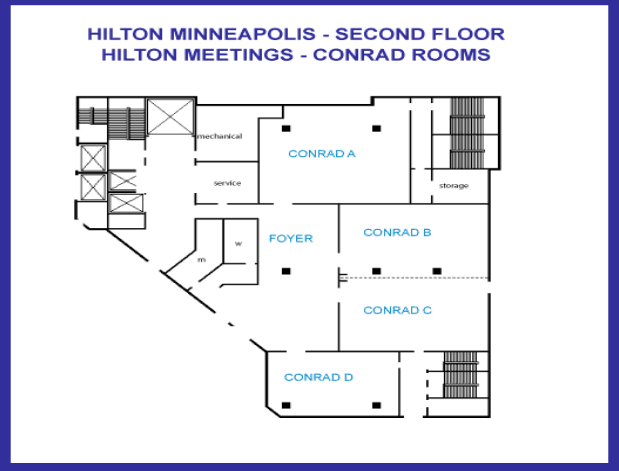
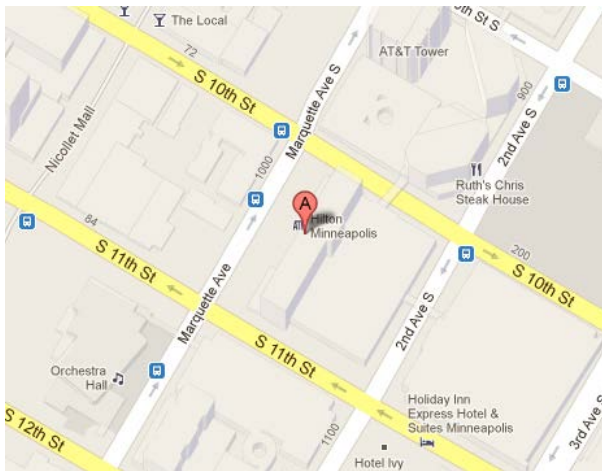
Hilton Minneapolis Hotel

The conference will be held at the Hilton Minneapolis Hotel. The address is: 1001 Marquette Avenue South, Minneapolis, MN 55403-2440, USA. Tel: 1-612-376-1000. See map on the next page.

Other Details

For other details not listed above, please see <http://www.scip.co/meeting-faqs/>

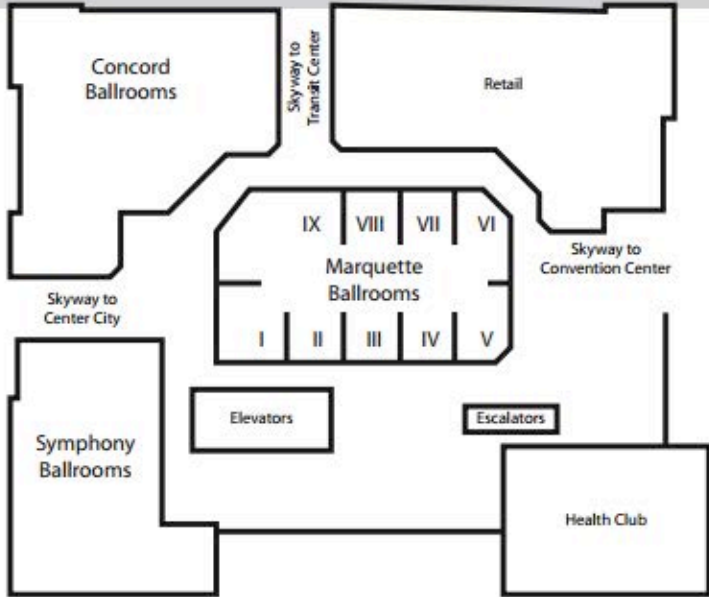
Below are the street map of the Hotel, floor plan of the Conrad Rooms (SCiP Registration and Spoken Sessions), Second Floor and the Convention Center (SCiP Poster Session)



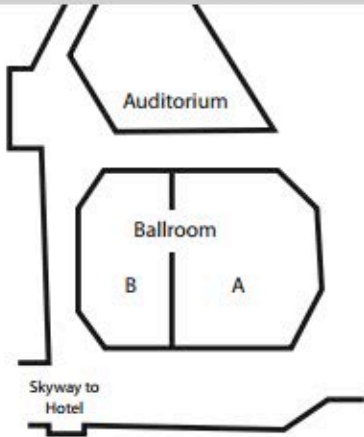
HILTON MINNEAPOLIS HOTEL & MINNEAPOLIS CONVENTION CENTER

- Keynote Address, Spoken Sessions, Other Meetings at the Hilton Minneapolis Hotel.
- Posters at the Minneapolis Convention Center - Level One.
- SCiP Annual Meeting at the Hilton Minneapolis Hotel, Conrad A, B, and C.

HILTON MINNEAPOLIS HOTEL - SECOND FLOOR



MINNEAPOLIS CONVENTION CENTER - LEVEL ONE



7:30 - 8:30am: Registration in Hilton Minneapolis Hotel (2nd fl., Conrad Ballrooms Foyer)

Conrad A	Conrad B
Session A: New Developments in Web-Based Data Collection <i>Chair: Ulf-Dietrich Reips</i>	Session B: Educational Tools <i>Chair: Ping Li</i>
8:15 Guillory, Hancock, Birnholtz, Gay, & Pollack	8:15 Widmer, Wolfe, Reyna, Cedillos, Fisher, & Brust-Renck
8:30 Klein, López-de-Ipiña, & Reips	8:30 Fisher, Wolfe, Reyna, Widmer, Cedillos, & Brust-Renck
8:45 Garaizar & Reips	8:45 Beringer
9:00 Blumer & Reips	9:00 Hwang, Hong, Liu, Cheng, & Lin
9:15 Reips	9:15 Tsai, Chen, & Lai

9:30 - 9:45: Coffee Break

Conrad A	Conrad B
Session C: Research Tools <i>Chair: Xiaowei Zhao</i>	Session D: Computational Linguistics and Corpus Data <i>Chair: Rick Dale</i>
9:45 Mueller & Piper	9:45 Hu, Graesser, & Cai
10:00 Krause & Lindemann	10:00 Hong, Hwang, & Tai
10:15 Zhang, Tsai, & Li	10:15 Magliano, Holt, & Durik
10:30 Zhao, Mauer, & Doyle-Smith	10:30 Titone
10:45 Parada & Busey	10:45 Paxton & Dale

11:00 - 12:00: Session E (Conrad C)

Keynote Address: Brian MacWhinney

Using the Web to study real-life language learning: measures, experiments, models, tutors, and tours

12:00 - 1:00: Lunch Break + Poster Setup

1:00 - 2:30: Session F: Poster Presentations (Convention Center Ballroom A)

2:30 - 3:45: Session G (Conrad C)

Presidential Symposium: Ping Li

Invited Speakers: Xiangen Hu, Charles Perfetti, & Chen Yu

Innovative Technologies and the Science of Learning

3:45 - 4:00: Coffee Break

Conrad A	Conrad B
Session H: Experimental Design and Analysis <i>Chair: Kay Livesay</i>	Session I: Cognitive Models and Networks <i>Chair: Michael Jones</i>
4:00 Anderson	4:00 Washburn, Beran, & Evans
4:15 Dixon	4:15 Epstein & Kirkish
4:30 Wang, Vaidyanathan, Haake, & Pelz	4:30 Li, Graesser, Cai, & Shaffer
4:45 Francis	4:45 Jones, Kievit-Kylar, Jacobsson, Tamir, & Siadati

Business meeting: 5:00 - 5:30 (Conrad C)

1:00-2:30pm: Poster Presentations (Level 1, Ballroom A; Minneapolis Convention Center)

Poster Presenters: Please follow the number designation below for your poster setup
(Note: Posters must be mounted between 12:00 - 1:00, and must be dismantled before 3:00pm)

1. Robidoux & Pritchard: Hierarchical clustering and nonword reading: Evaluating computational models of reading
2. Tang, Cai, & Hu: A measure for word similarity in LSA
3. Rolotti, Zinszer, & Li: Bayesian inference for bilingual word learning
4. Broussard & Livesay: The influence of social priming on speech perception
5. Curtis & Jamieson: An instance-based account of mental arithmetic performance
6. Jamieson, Crump, & Hannah: A computer tool to simulate and teach associative learning
7. Zhang & Anderson: Investigating the dynamics of search and decision making
8. Miclat, Burgess, & Estep: Modeling the concept of “obscurity”
9. Li, Graesser, & Cai: Formality prediction with data mining: Chinese political leaders as a case study
10. Viray, Wong, Chen, Burgess, & Maples: Normative results for gender usage of emoticons: Reliability of meaning
11. Washburn, Schultz, Gullledge, & Phillips: Comparing fTMS and fTCD for studies of functional cerebral asymmetries
12. Chan, Chen, Chang, & Chang: Garden path sentences between Han and Indigene: an fMRI study
13. Chen, Cha, Chang, Sung, & Hsieh: CRIE: A tool for analyzing Chinese text characteristics
14. Chang, Sung, Lee, & Hsieh: A Chinese word segmentation and POS tagging system for readability research
15. Hong, Tseng, Li, & Sung: Constructing a Chinese text readability formula with multi-level linguistic features
16. Lin, Chen, & Sung: Integrating Cloze tests with Item Response Theory to validate a readability formulae
17. Tseng, Chang, & Sung: Evaluation of the feasibility of online readability application
18. Wolfe, Reyna, Cedillos, Widmer, Fisher, & Brust-Renck: Helping women decide about testing for genetic breast cancer risk with an Intelligent Tutoring System
19. Williams & Hu: Exploring the effectiveness of a novel feedback mechanism within an intelligent tutoring system
20. Collins & Freeman: Training the cognitive benefits of video games: Can massively multiplayer online role playing games also improve performance?
21. Koch & Koch: Shyness, positive and negative affect, life satisfaction and video games
22. Beringer: End user development in psychology teaching
23. Paniukov, Arslan-Ari, & Crooks: Examining the modality effect in multimedia self-paced instruction
24. Arslan-Ari, Paniukov, & Crooks: A redundancy effect in self-paced multimedia instruction
25. Liao, Huang, Xie, & Hu: Unified data repository for psychology experiments
26. Acha, Salaburu, & Laka: EHME: A new word database for research in Basque language
27. Stoet: PsyToolkit: Using the web to program and share experiments
28. Garaizar, Perez-Cubillas, & Matute: Libet’s clock goes online: Using Labclock and Labclock Web to conduct offline and online studies on conscious will
29. Koch & Hotovec: An Android app for the Nonverbal Stroop Card Sorting Task
30. Gomez & Zimmerman: S-Delta plots; a method to display RT distributions across multiple subjects
31. Ohyanagi, Kanaya, Sengoku, & Miyazaki: Preliminary results of new reaction time tasks of detecting velocity change of a moving stimulus to assess visual attention skills in Occupational Therapy
32. Plant & Turner: Could your equipment account for your experimental effect?

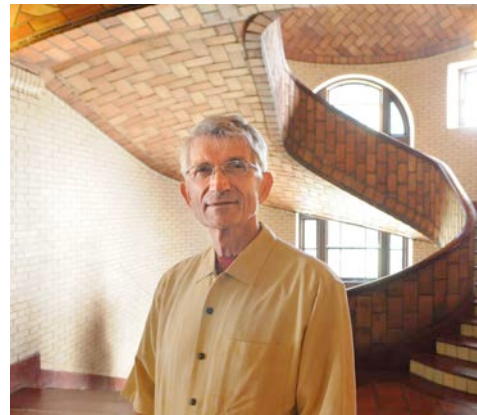
Keynote Speech

Using the Web to study real-life language learning: Measures, experiments, models, tutors, and tours

Brian MacWhinney

Carnegie Mellon University

Understanding the process of second language acquisition (SLA) is a fundamental goal for psychology and linguistics. However, SLA research has been unable to develop a methodology that permits the in vivo testing of empirical hypotheses. Classroom studies have inadequate control of conditions, and experimental studies fail to align with the dynamics of actual language learning. Fortunately, the advent of mobile computing devices (tablets and phones), ubiquitous 3G/4G networks, HTML5, and methods for web-based data collection provides us with a unique and exciting opportunity to advance the scientific basis of SLA theory and the practice of second language instruction.



This talk explains the design of a system called the Language Partner that combines practice in basic skills such as vocabulary and grammar with contextualized practice in real language use. The system is designed for learners who are immersed in a second language environment while attending university language classes in cities like Beijing, Granada, or Lyon. All facilities are configured to run on a tablet or portable computer with 3G/4G connections. Data are logged continuously to servers at CMU for analysis through predictive data-mining methods using HMM student models. Materials also include experiments that are designed into the instructional materials.

The Language Partner uses a series of on-line individual difference measures of memory, attention, and learning to assess student abilities and preferences. Contextualized learning is promoted by allowing instructors to configure online exercises to match the textbook and the current lesson plan. Learners are also given access to specific embedded tutorials within Google Maps Tours for interacting with bus drivers, salespeople, and officials in restaurants, museums, and elsewhere in the city. In addition, the Language Partner provides methods for easy vocabulary access, translation, picture recognition, and text-to-speech production. Data collected from the Language Partner can be used to test alternative SLA theories and to measure ways in which alternative instructional treatments can optimize learning.

Presidential Symposium

Innovative technologies and the science of learning: Language acquisition
by children and adults

Ping Li

Pennsylvania State University

Computational Linguistics Applications in Intelligent Tutoring Environment

Xiangan Hu

University of Memphis

Computational Linguistics (CL) is one of the most important enabling theories and technologies used at the Advanced Distributed Learning Center for Intelligent Tutoring Systems Research and Development (ADL C ITS R&D). In this talk, we will present some recent research and development efforts on computational language studies. Specifically, we will present 1) Theory and implementation of semantic spaces (SS), 2) Text cohesion analysis based on suite of CL utilities, and 3) Application of CL in advanced learning environments.

Understanding Micro-level Multimodal Behaviors in Early Word Learning and Social Interaction

Chen Yu

Indiana University

I will present a set of studies using advanced sensing equipment and data mining techniques to collect and analyze dynamic multi-coupling sensory data in child-parent interactions. The results show that inter-agent coordination at least at an elementary level depends crucially on external (and observable) behaviors by the participants where the real-time behavior of one participant organizes the actions of the other – behaviors such as eye movements, head turns, and hand gestures – and those micro-level behaviors are predictive to successful word learning through child-parent free-play interaction.

LearnLab: A technological-social Infrastructure for the learning sciences

Charles Perfetti & Ken Koedinger

University of Pittsburgh and Carnegie Mellon University

Scientifically rigorous research can be carried out in educational settings. The NSF-funded Pittsburgh Science of Learning Center (LearnLab) has developed a theoretically-focused and technology-supported program of such research. I review the important role that technology plays in this effort, especially the storage of large and very fine-grain learning data (LearnLab's DataShop), associated tools to support data mining and model building, and user-friendly software for the development of cognitive tutors. The technology supports the learning sciences through their value for hypothesis testing and theory building and because of their availability to researchers around the world.

8:15-9:30 Session A

New Developments in Web-Based Data Collection

8:15 Novel ecological momentary assessment tools for studying deception and health behavior

Jamie Guillory ¹, Jeffrey Hancock ¹, Jeremy Birnholtz ², Geri Gay ¹, J. P. Pollack ¹
Cornell University, USA ¹, Northwestern University, USA ²
jeg258@cornell.edu, jth34@cornell.edu, jpb277@cornell.edu, gkg1@cornell.edu,
jpp9@cornell.edu

Self-report methods for assessing attitudes and behavior are criticized for being victim to recall bias and reporting error, but also because they require people to consider their current mental state in addition to contextual and cultural factors related to the measurement. Ecological momentary assessment (EMA) is a class of self-report data collection methods that remedies these shortcomings by assessing attitudes and behavior in situ using computers or mobile devices. We developed EMA tools to study deception and health behavior. In our deception research, we developed and implemented a method that required participants to assess deceptiveness of instant messages they sent over a 4-day period. To study health, we developed a mobile application that participants used to photograph, comment and rate healthiness of daily health behaviors. These novel methods provide a rich representation of the context and motivations behind deception and health behaviors.

8:30 Assessing mobile peer-to-peer micro-services from a user-centric and large scale community perspective in a living lab scenario

Bernhard Klein ¹, Diego López-de-Ipiña ¹, Ulf-Dietrich Reips ^{1,2}
University of Deusto, Spain ¹, IKERBASQUE, Basque Foundation for Science, Spain ²
u.reips@ikerbasque.org

We report on innovative social location-aware services for mobile phones. With these micro services users can tag their environment with textual comments and photos and share these with others. To gain a better understanding of user acceptance, empirical frameworks like the Compass Acceptance Model consider factors of user-interface, technology, community, and environment. Even though such frameworks are able to separate short-term from long-term user benefits, they fail in providing detailed information about characteristic usage patterns. For this reason we implemented the Living Lab that emulates a realistic ecosystem in an urban context. By installing a data logger application on the phone, we can record service usage data, current context, invoked services, and content provided by the user. Service usage is analyzed from the user centric and community perspective. Living Lab was shown to identify usability problems, gain data on characteristic usage patterns and hotspots and to evaluate a service's effectiveness.

8:45 **Social Lab: A social engineering wargame**

Pablo Garaizar ¹, Ulf-Dietrich Reips ²

Universidad de Deusto, Bilbao, Spain¹, University of Deusto, Spain, IKERBASQUE, Basque Foundation for Science ²

garaizar@deusto.es, u.reips@ikerbasque.org

Social networking has surpassed email and instant messaging as the dominant form of online communication (Meeker, Devitt & Wu, 2010). Users of social networks often overlook dangerous implications of settings of privacy configurations. Although these may cause unforeseen problems, it is difficult to educate users about privacy. Social Lab faces this problem by providing a game-based learning environment. In Social Lab users play the role of someone trying to invade another's privacy. They have to solve automated social engineering challenges, e.g., the player has to post on the wall of a friend of the victim/bot to look friendly and be accepted as a friend. Because daily social networking situations are simulated, Social Lab can be the way to make users aware of their own privacy setting and the dangers that lie in them.

9:00 **Testor: A Web-based platform for the development and management of psychological tests**

Thomas Blumer ¹, Ulf-Dietrich Reips ²

University of Zurich, Switzerland ¹, University of Deusto, Spain, IKERBASQUE, Basque Foundation for Science ²

thomas.blumer@psychologie.uzh.ch, reips@deusto.es

Internet-based data collection has many advantages in various fields. In psychological testing, the Internet enables researchers to efficiently collect and publish royalty free items and scales, which are essential for the development of "open source" tests. To support this development, we present Testor, a Web-based platform for the development and management of psychological tests. Testor serves as a repository by providing facilities for sharing items and scales. At the same time, Testor enables researchers to develop and administer Internet-based tests without requiring programming skills, thereby allowing them to utilize advantages of online data collection like easy access to large samples in order to facilitate test validation. Furthermore, the application implements requirements based on empirical findings and recommendations from the International Test Commission in order to foster best practices for online testing and prevent users from common pitfalls. Finally, we present future plans of development. Testor is available at <http://testor.rgbulls.com/>.

9:15 **Big Data: Using Google services for research**

Ulf-Dietrich Reips

University of Deusto, Spain, IKERBASQUE, Basque Foundation for Science

u.reips@ikerbasque.org

Large collections of entries or traces from human behavior on the Internet ("Big Data") have become an accessible source for research. Examples include the definition of points of interest via data mining in uploaded pictures (Barras, 2009), prediction of influenza outbreaks from searches (Ginsberg et al., 2009), and our own work on attributions of personality characteristics to first names accessed via Twitter mining (Reips & Garaizar, 2011). Recently Google has created freely available interfaces to their search data. These include Google Trends, Google Insights, Google Correlations, and more specific services like Google Flu Trends and Google Eurovision. I show how these services can be used in psychological research, sometimes with profound results achieved within minutes, and where current limitations are. I will discuss ethical issues and how science principles are sometimes at odds with characteristics of Big Data services provided by companies.

8:15-9:30 Session B

Educational Tools

8:15 Let's talk about breast cancer: An analysis of tutorial dialogues between women and an Intelligent Tutoring System

Colin Widmer¹, Christopher Wolfe¹, Valerie Reyna², Elizabeth Cedillos¹, Christopher Fisher¹, Priscila Brust-Renck²
Miami University¹, Cornell University²
widmercl@muohio.edu, wolfecr@muohio.edu, vr53@cornell.edu, cedillem@muohio.edu,
fisherc2@muohio.edu, pgb56@cornell.edu

We used AutoTutor Lite (ATL) to create an Intelligent Tutoring System (ITS) guided by Fuzzy-Trace Theory to teach women about genetic breast cancer risk. Sixty-five women interacted with ATL including five interactive tutorial dialogues where avatars asked questions and responded to participants' verbal input. Tutorial dialogues centered on questions including, "How do genes affect breast cancer risk?" and "What is the case for genetic testing for breast cancer risk?" ATL "understands" participants using Latent Semantic Analysis (LSA) and "expectations texts." We developed a simulation technique to set LSA parameters and create expectations texts. We analyzed tutorial interactions using reliable scoring rubrics for content; assessments of ATL response quality; and ATL coverage scores. Participants covered much of the content. ATL generally responded appropriately, and appropriate responding yielded higher participant knowledge scores. Tutorial dialogues are partially responsible for ATL efficacy. Having ITSs assist argumentation in ill-structured knowledge domains is novel and promising.

8:30 **Evaluating gist-based discrimination of genetic breast cancer risk using multiple signal detection**

Christopher Fisher ¹, Christopher Wolfe ¹, Valerie Reyna ², Colin Widmer ¹, Elizabeth Cedillos ¹, Priscila Brust-Renck ²
Miami University ¹, Cornell University ²
fisherc2@muohio.edu, wolfecr@muohio.edu, vr53@cornell.edu, widmercl@muohio.edu,
cedillem@muohio.edu, pgb56@cornell.edu

We develop an instrument using Multiple Signal Detection (MSD) measuring how people assess genetic breast cancer risk. We created 12 scenarios describing women of low, medium, and high risk based on the Pedigree Assessment Tool and Gail Model. Participants made ordinal "gist" assessments of low (L), medium (M), or high (H) risk. The MSD model quantifies discrimination among risk categories (d') and provides three discriminability measures distinguishing among risk levels: $d'(L-M)$; $d'(M-H)$; and $d'(L-H)$. In Study 1, untrained participants modestly distinguished among breast cancer risk levels. In Study 2, participants were randomly assigned to either Intelligent Tutoring System (ITS); National Cancer Institute website; or control groups before rating scenarios. Nonparametric permutation tests revealed significantly increased discriminability across all risk categories for the ITS compared to the control group, $p < 0.0005$, without affecting the response bias. This easy to administer technique is powerful and theoretically compatible with ordinal gist-based discrimination judgment tasks.

8:45 **On-Demand Cognition Library**

Joerg Beringer
BeriSoft Cooperation
joerg.beringer@berisoft.com

The On-Demand Cognition Library is a new cloud-based offering to help teaching staff to enrich their psychology classes with hands-on exercises that let students directly experience experimental paradigms and collect data. The concept of web-based experiments is not new, but the offerings are mostly fragmented originating from individual investments or are add-ons of on-premise software packages. The Cognition Library is a next generation solution that builds on the ERTS on-demand platform to fully leverage modern web technology and the increasing runtime accuracy as well as graphics capability of standard web browsers. A truly zero-footprint deployment opens up new opportunities of reaching students at home and scaling the number of test stations by leveraging personally owned computers. Viral spreading within social networks can become a key mechanism on recruiting subjects in context of educational exercises or research that targets populations other than psychology students.

9:00 **Using gesture-based interactive games to improve Chinese pronunciation for native-like immigrated students**

Ming-Yueh Hwang, Jon-Chao Hong, Yeu-Ting Liu, Hao-Yueh Cheng, Pei-Shin Lin
National Taiwan Normal University
t06013@ntnu.edu.tw, tcdahong@gmail.com, yuetingliu@ntnu.edu.tw, haoyueh@ntnu.edu.tw,
willcc107@hotmail.com

In these days, there are more and more immigrant students have come into elementary schools in Taiwan. For these students, Chinese pronunciation may affect the learning of them. This study designed a gesture-based interactive game (GBIG) for teachers to incorporate into inclusive education programs (IEP). To examine the effectiveness of GBIG in IEP, this study invited 108 3rd and 4th grade students to practice. The results indicated that language experience had a negative association with language anxiety, but language anxiety had an effect on participants' interest in playing GBIG, which was mediated by cognitive load. This implied that the GBIG would be beneficial to students with low language anxiety by increasing their mental effort and interest in the IEP.

9:15 **The effect of learning Chinese language and culture on children's creativity by using APP of digital archives of the National Palace Museum**

Yahsun Tsai ¹, Peng-Fei Chen ², Ting-Sheng Lai ³
Department of Applied Chinese Language and Culture, National Taiwan Normal University ¹,
National Taiwan Normal University ², National Palace Museum ³
yahsun@ntnu.edu.tw, blissfulalice@ntnu.edu.tw, samlai@ms96.url.com.tw

This study is conducted to examine the impact of learning Chinese language and culture on children's creativity by applying APP concerned to digital archives of the National Palace Museum. 90 American students (aged 10 -12) are assigned randomly and equivalently into one experimental and two control groups. Two different interventions based on resource of the National Palace Museum implement to the experimental group and control group A over 8 lessons. Non-treatment is set to control group B. The students in the experimental group A learn Chinese language culture by assistance of use APP. In contrast, the Chinese lessons in control group B do not have the treatment of using APP. Creativity will be evaluated by consensual assessment of 3 raters and deterring the statistical results.

9:45-10:45 Session C

Research Tools

9:45 **Ten years of open source psychology: The Psychology Experiment Building Language and Test Battery**

Shane Mueller ¹, Brian Piper ²
Michigan Technological University ¹, Husson University ²
shanem@mtu.edu, psy391@gmail.com

The Psychology Experiment Building Language (PEBL) is approaching the ten-year anniversary of its initial development and release. In that time, it has been downloaded more than 100,000 times, and its test battery has grown to more than 60 tests, all of which are free to use, modify, and share with other researchers. We will discuss the current state of the PEBL system, review some of the published research that has used PEBL tests, and discuss the benefits and challenges of open-source development for scientific software.

10:00 **Expyriment: A light-weight Python library for the script-based development of experiments**

Florian Krause ¹, Oliver Lindemann ²
Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands ¹, Division of Cognitive Science, University of Potsdam, Germany ²
f.krause@donders.ru.nl, oliver.lindemann@uni-potsdam.de

Expyriment is an open-source and platform independent light-weight Python library for designing and conducting timing-critical behavioural and neuroimaging experiments. It has been tested extensively under Linux and Windows. In contrast to other packages, Expyriment focuses on the power and flexibility of programming languages and does therefore not aim to be an integrated environment with a graphical use interface. The major goal is instead to provide a well-structured Python library for a script-based experiment development with a high priority on the readability of the resulting programme code. Expyriment is an all-in-one solution, as it handles the stimulus presentation, recording of I/O events, communication with other devices and the collection and preprocessing of data. It offers furthermore a hierarchical design structure, which allows an intuitive transition from the experimental design to a running programme. Expyriment is therefore also suited for students as well as experimental psychologists and neuroscientists with little programming experience.

10:15 **A modularized, dynamic, cloud-based Language History Questionnaire (LHQ)**

Fan Zhang, Erlfang Tsai, Ping Li
Pennsylvania State University
fvz5016@psu.edu, ext11@scasd.org, pul8@psu.edu

Language history questionnaire (LHQ) is an important tool for assessing language learners' linguistic background, the context and habits of language use, proficiency in multiple languages, and the dominance and cultural identity of the languages acquired. Outcomes from such assessments have often been used to predict or correlate with learners' linguistic performance in cognitive and behavioral tests. Previously we identified the most commonly asked questions in published questionnaires and proposed a generic LHQ (Li, Sepanski, & Zhao, 2006). Taking advantage of the dynamic features of web-based interfaces, we have implemented a new cloud-based LHQ, in four different modules to suit different researchers' focuses and needs (history, usage, proficiency, and dominance). The new LHQ will allow investigators to dynamically produce their own LHQ on the fly, and allow participants to complete the LHQ online through individualized URLs. The results are saved in a spreadsheet for all participants who have completed the LHQ, eliminating the needs of manual coding of LHQ results. The investigators can view, download, and delete the LHQ results on the web. Privacy issues are handled through online assignments of ID numbers for experiments and recording of data with only participant numbers.

10:30 **A general music background questionnaire based on Google Forms and Google Template**

Xiaowei Zhao, Meaghan Mauer, Noah Doyle-Smith
Emmanuel College
zhaox@emmanuel.edu, mauerm@emmanuel.edu, doylesmithn@emmanuel.edu

The influence of music training on people's cognitive abilities has become an important topic in psychology. In this study, we have examined twelve published music background questionnaires and identified the most commonly asked questions on evaluating participants' music background and training history. Based on these general questions, we developed the ECMBQ (Emmanuel College Music Background Questionnaire). An online interface of the questionnaire was created based on the technology of Google Forms, through which we can collect and save participants' responses in a spreadsheet for further analysis. In addition, to facilitate the ease of data collection of other researchers in the field, we transferred the questionnaire into a Google Template. Any colleague who has a Google account can easily apply our template into her website, make any necessary changes for her research purpose and collect responses from participants for her own use. A website of this questionnaire can be found at <https://sites.google.com/site/ecmbqzhao1/>.

10:45 **ExpertEyes: Open source eyetracking software and hardware for research and teaching**

Francisco J. Parada, T. A. Busey
Indiana University
fjparada@indian.edu, busey@indiana.edu

ExpertEyes is a low-cost, open source package designed to provide portable high-definition eye tracking. The project involves several technological innovations, including portability, high definition video recording, and multi-platform software support. It was designed for challenging recording environments and all processing is done offline to allow for optimization of parameter (re-)estimation, making it suitable for scientific research and teaching purposes among other uses. Standard drift correction procedures are included, as is support for export to Matlab and similar processing environments. The accuracy of the system is comparable to commercial eyetracking systems, with a typical accuracy of less than .5 degrees and best accuracy below 0.24 degrees. Our software is freely downloadable along with complete hardware plans. In this symposium we will go over our software and our suggested high-definition hardware. Finally, we will discuss software extensions and provide examples of usage in different areas of scientific research.

9:45-10:45 Session D

Computational Linguistics and Corpus Data

9:45 **Analysis of social media: Potential application of individualized domain-specific and context sensitive semantic analysis**

Xiangen Hu, Art Graesser, Zhiqiang Cai
The University of Memphis
xhu@memphis.edu, art.graesser@gmail.com, zhiqiang.cai@gmail.com

We first present a simple model for social media. This model describes how social media such as twitter, facebook, or corporate internal messaging system work. The basic components of this model are memes and agents. The model is created based on the idea of a popular book “virus of the mind” (Richard Brodie, 2008). In this model, Memes are implicit semantic dimensions contained in text messages and agents are senders and receivers of messages. We explain how individualized domain-specific context-sensitive semantic analysis may help us analyze social media.

10:00 **Usability affects mobile English vocabulary learning**

Jon-Chao Hong, Ming-Yueh Hwang, Kai-Hsin Tai
National Taiwan Normal University
tcдахong@gmail.com, t06013@ntnu.edu.tw, star99xin@gmail.com

Today, rapid development of mobile technology is increasing the integration of mobile devices and subject learning has created greater flexibility for students to participate in learning. The present study developed a systematized way for learning 1,000 fundamental English vocabulary words via self-assessment as an iPhone App as “M-Learning” based of Fibonacci Sequences, called English Vocabulary Learning @ Star (EVL@S) to motivate learners to learn at their convenience. Research data from 107 participants was collected out of 10,000 registered users. The result shows that the usability of the learning system may not affect users’ motivation by using the system to learn English.

10:15 Assessing the dynamic nature of comprehension processes

Joe Magliano ¹, Janet Holt ², Amanda Durik ¹
Northern Illinois University ¹, Illinois Education Research Council ²
jmagliano@niu.edu, jholt@niu.edu, adurik@niu.edu

The current study investigates how comprehension processes dynamically occur over the course of reading a text. The goals of the current study were to 1) explore the dynamic pattern of comprehension processes, as measured by elaboration, bridging, and paraphrasing processes that readers demonstrate when producing concurrent verbal protocols and how these are manifested over the course of text, and 2) to determine how these patterns are different based on reading skill level and text specific prior knowledge. Participants produced concurrent verbal protocols while reading expository texts in the context of the Reading Strategy Assessment Tool, which automatically score the protocols in terms of evidence for bridging, paraphrasing, and elaboration. A novel application of growth curve modeling was used to assess how these processes change over the course of the texts. The results are interpreted in terms of the stages of structure building during reading.

10:30 The psychological reality of corpus derived measures of idioms

Debra Titone
Department of Psychology, McGill University
dtitone@psych.mcgill.ca

Idioms are multiword sequences that reliably bridge between linguistic form and meaning (kick the bucket). Psycholinguistic studies rely on human processing measures, whereas applied linguistics or usage-based studies rely on distributional patterns in large-scale corpora. We tested whether basic corpus measures relate to human performance for 60 English idioms (taken from Libben & Titone, 2008) using COCA (Davies, 2008). Increased mutual information correlated positively with final-word predictability, negatively with component word frequency, and played a prominent role during L2 but not L1 reading, suggesting a link to lexicalization. In contrast, the increased ratio of figurative to non-figurative usage correlated positively with decomposability, and played a prominent role during both L1 and L2 reading (Titone & Lovseth, in revision), particularly for later integration stages. Thus, corpus-derived measures of idioms differentially constrain how we read

and make decisions about idioms, presumably because they reflect how we experience and use idioms normally.

10:45 **Linguistic alignment in debate**

Alexandra Paxton, Rick Dale
University of California, Merced
paxton.alexandra@gmail.com, rdale@ucmerced.edu

Debate has been considered a standard practice of academia for centuries. However, it has not received much attention from experimental psychology, even to characterize it as a mode of communication. Our previous work in interpersonal alignment has found differences in bodily movement synchrony during argument: it is, on average, hindered (Paxton & Dale, in preparation). In the current project, we extend these findings about interpersonal alignment in asymmetric interactions to debate: We explore the linguistic alignment among individuals engaged in debate, as evidenced through transcript analysis. Transcripts were taken from publicly available websites that involved debate on hot-button political, social, and cultural topics. In our presentation, we will showcase dynamic discourse analysis methods that reveal distinctions in interpersonal alignment patterns between debate opponents. Results further suggest that argument is a distinct conversational mode that modulates the dynamics of language usage.

4:00-4:45 Session H

Experimental Design and Analysis

4:00 **Permutation tests of null hypotheses are truly non-parametric for experimental data – and when the data are dichotomous, for non-experimental data**

Richard Anderson
Bowling Green State University
randers@bgsu.edu

Permutation tests have been promoted as a way to test null hypotheses without having to make distributional assumptions. Nevertheless there have been concerns that permutation tests of differences between means may entail an equal-variance assumption. The present paper clarifies that permutation tests assume only exchangeability-under-the-null, thus permitting the researcher to test the very specific null hypothesis that the coupling of the data points to group labels has occurred by chance. The paper also emphasizes that this sole criterion, exchangeability-under-the-null, is always met in proper, true experiments - and when the data are dichotomous, in non-experiments. Finally the paper describes the conducting of permutation tests on entire populations (rather than samples) to highlight that the fact that the meaning of the null hypothesis, and therefore of the p value, differs for parametric versus permutation tests. Thus one type of p value is

not a standard for evaluating the validity of the other.

4:15 **The effective number of parameters in post hoc models**

Peter Dixon
University of Alberta
peter.dixon@ualberta.ca

Although a parsimonious, post hoc model may involve relatively few parameters, there may be many other potential model parameters that were, in effect, fixed at 0 after examining the data. Because these fixed parameters could have been varied if the data had come out differently, one could argue that they should be included when assessing the complexity of the model. On the other hand, intuitively, it seems as if there should be some advantage that accrues to identifying a simple and compelling description of the data, even if it is post hoc. This problem is considered in the context of factorial designs in which a potentially parsimonious description of the results consists of a limited set of simple effects. Monte Carlo simulations are used to establish the effective number of parameters for various classes of such simple, but post hoc, models.

4:30 **Are eye trackers always as accurate as we assume?**

Dong Wang¹, Preethi Vaidyanathan¹, Anne Haake², Jeff Pelz¹
Chester F. Carlson Center for Imaging Science, College of Science, Rochester Institute of Technology¹, Golisano College of Computing and Information Sciences, Rochester Institute of Technology²
dxw1481@rit.edu, pxv1621@rit.edu, anne.haake@rit.edu, jeff.pelz@rit.edu

The accuracy of eye trackers is vital in eye-tracking research, but there is no established international standard to measure their accuracy. Many practitioners simply accept the manufacturer's stated accuracy. We did a pilot study to measure the accuracy of a commercial eye tracker during an eye-tracking experiment by inserting a test slide between every two stimuli. The test slide had a small, visible target within a trigger area of interest (tAOI) that was invisible to the participant. The trial continued when the participant's gaze stayed within the tAOI for 150ms. Measured accuracy was defined as the distance between the target and the triggering gaze position within the tAOI. That distance exceeded the calibration error on about half the trials, and average measured accuracy was 45% worse than the manufacturer's stated accuracy. The methodology we introduced can better characterize the actual uncertainty of eye-tracking data and help researchers better understand observers' viewing patterns.

4:45 **It doesn't replicate, great!: Reasons for a failure to replicate in experimental psychology**

Gregory Francis
Purdue University
gfrancis@purdue.edu

Experimentalists emphasize successful replication. When a finding is not replicated there is often a vigorous debate about methods, stimuli, calibration, instructions, and participant characteristics.

Sometimes the debates become quite heated, with accusations of incompetence; and many observers closely watch these debates to learn who is ultimately correct. These debates reflect misunderstandings about the properties of hypothesis testing and the nature of empirical studies. Given the modest power of most studies in experimental psychology findings should not always successfully replicate (e.g., reject the null hypothesis). Random sampling insures that researchers will sometimes draw a sample that does not show the effect, even if it exists in the population. A set of experiments that always found evidence of an effect would be unbelievable, so reported failures to replicate can increase confidence that the experimental findings are valid. The properties of replication provide a new interpretation of some recent debates in experimental psychology.

4:00-4:45 Session I

Cognitive Models and Networks

4:00 How 25 Years with the “Rumbaughx” has changed comparative cognition

David A. Washburn, Michael J. Beran, Theodore Evans

Georgia State University

dwashburn@gsu.edu, mjberan@yahoo.com, theodore.evans@gmail.com

About 25 years have passed since rhesus monkeys joined the ranks of computer users. Dedicated computerized test stations and game-like computerized tasks have now become standard testing apparatus for studies of comparative cognition with a wide range of primate species, but it bears noting that the paradigm was high-risk and innovative at the time. Because new psychological apparatus not only allow a host of new questions to be asked but also shape in some ways the answers that are produced, the field of comparative cognition has changed dramatically during this period, as has our understanding of the cognitive competencies of nonhuman primates. This paper will highlight these changes and the role that the “Rumbaughx” (in honor of Duane Rumbaugh, who developed the paradigm for monkeys, and to parallel other significant comparative apparatus like the puzzle box, the shuttle box, and the Skinner box) has played in the discipline.

4:15 Do humans know how humans differ from computers?

Robert Epstein ¹, Gina Kirkish ²

American Institute for Behavioral Research and Technology ¹, University of California, San Diego ²
re@aibr.org, gkirkish@ucsd.edu

In an Internet-based study with an ethnically-diverse group of 925 subjects in the US, Canada, and 50 other countries (mean age 25.3, SD=12.4), people were asked questions that revealed whether they could identify subtle characteristics of human behavior and emotion which, at the present time, distinguish humans from computers. To our surprise, many subjects were poor at this task,

even when asked fairly simple questions about human relationships, scoring 57 percent on average on our 66-item questionnaire. Sensitivity to what it means to be human was a good predictor of whether someone was employed and a modest predictor of other self-reported outcomes, including personal and professional success. Scores increased with age, and females outscored males by a small margin. In general, the study suggests that as computers become more human-like, most people will have enormous difficulty distinguishing them from humans.

4:30 **Data mining analysis of online chats in an epistemic game**

Haiying Li¹, Arthur C. Graesser¹, Zhiqiang Cai¹, David W. Shaffer²
University of Memphis¹, University of Wisconsin-Madison²
haiyinglit@gmail.com, art.graesser@gmail.com, zhiqiang.cai@gmail.com,
david.williamson.shaffer@gmail.com

Land Science is an epistemic game in which players develop land use plans as interns at an urban and regional planning firm with the guidance of the mentor. At the first stage, the mentor is the human mentor. Then with the analysis of the conversations between players and human mentor, the auto-mentor will be developed. This paper uses data mining tool “WEKA” first to cluster the players’ and mentor’s conversations and then to find out the best prediction of the mentor’s conversational patterns with the features of speech act, state transition network, epistemic network analysis, newness and relevance along with the game elements including time slots, groups, and tasks. The benefits of data mining are its ability to gain deeper, potential patterns of the multi-party online chats, and provide the opportunity to develop production rules for the conversations between players and the auto-mentor in the epistemic game system.

4:45 **Using cognitive models to identify efficient, secure, and memorable passwords**

Michael Jones¹, Brent Kievit-Kylar¹, Markus Jakobsson², Dahn Tamir³, Hossein Siadati²
Indiana University¹, PayPal Inc.², Knewton Inc.³
jonesmn@indiana.edu, bkievitk@indiana.edu, markus.jakobsson@gmail.com, dtamir@gmail.com,
s.h.siadaty@gmail.com

We describe a new framework for generating robust and memorable passwords based on computational models of human memory. The framework takes advantage of spellcheck systems and uses only ‘real’ words, resulting in faster entry time on mobile devices, but while maintaining high resilience to guessing. The user generates a sequence of words given a probe to describe an episodic memory. To authenticate, the user is probed with the original word, cuing the episodic trace. The likelihood of guessing the response words by an individual who does not have the encoded individual memory is low, but the cue is very effective at producing retrieval by the individual who generated it, acting as a sort of mnemonic RSA key. We explore lexical characteristics of cue words that produce password sequences that are both easy to recall and low in guessability, and present a model to allow fuzzy authentication based on semantic space models.

1:00-2:30 Session F

Poster Session

1. **Hierarchical clustering and nonword reading: Evaluating computational models of reading**

Serje Robidoux, Stephen Pritchard

ARC Centre of Excellence for Cognition and its Disorders, Macquarie University

serje.robidoux@mq.edu.au, stephen.pritchard@mq.edu.au

Pritchard, Coltheart, Palethorpe, and Castles (2012) reported nonword reading aloud data comparing responses from CDP+/CDP++ (Perry, Ziegler, & Zorzi, 2007; 2010), DRC (Coltheart et al., 2001), and 45 individual subjects. They reported simple analyses based on frequency of responses and types of errors, and found that DRC out-performs CDP+/++ in nonword reading. That approach leaves an important question unanswered: could it be that the models are describing different classes of subject? Hierarchical clustering techniques applied to the model and subject response data sheds some light on this question.

2. **A measure for word similarity in LSA**

Quan Tang, Zhiqiang Cai, Xiangen Hu

University of Memphis

tangbaishui@gmail.com, zhiqiang.cai@gmail.com, xiangenhu@gmail.com

The theory and application of semantic spaces (SS) have been applied in natural language processing for various purposes, such as information retrieval, text analysis, and advanced learning environments. One of the challenges for research and development of SS is to compare SS based on different methods, domains, and context. Hu et al. (2005) provided a general framework of SS and method of measuring similarities between SS by introducing three levels of similarity measures between any two tokens from any two SS based on different methods, domains, and context. In this paper, we demonstrate the utility of one of the measures introduced by Hu et al. (2005). We will provide simple computation for this measure and provide a meaningful interpretation by examining a few different SS.

3. **Bayesian inference for bilingual word learning**

Sebastian Rolotti, Benjamin Zinszer, Ping Li

Pennsylvania State University

svr5177@psu.edu, bdz107@psu.edu, pul8@psu.edu

The indeterminacy problem describes the challenge that infants face in deciding which words refer to which objects in their environment. Bayesian models use probabilistic inferences to resolve this induction problem and show improved performance over other computational models in constructing potential lexicons and inferring speakers' referential intentions. In this study we investigate a Bayesian model's ability to learn in more complex situations, first with more objects than in previous research and then in a bilingual scenario where more than one word refers to the same object. We found that the model's absolute and relative performance was attenuated with increased complexity, but when the mutual exclusivity constraint was relaxed, the model's performance in bilingual word learning was improved.

4. **The influence of social priming on speech perception**

Sierra Broussard¹, Kay Livesay²
University of California, Irvine¹, Linfield College²
sbrouss@linfield.edu, kaylivesay@gmail.com

Speech perception relies on auditory, visual, and motor cues. Fuzzy Logic Model of Perception (FLMP) suggests that if one of these types of speech modes is altered, the perception of that speech signal should be altered in a quantifiable and predictable way. The current study uses social priming to activate the schema of blindness in order to reduce reliance on visual cues of syllables with a visually identical pair (e.g., /ba/, /pa/). According to the FLMP, by lowering reliance on visual cues, visual confusion should also be reduced, allowing the visually confusable syllables to be identified more quickly. Data revealed that some syllables (/ba/) showed reliable facilitation, while other syllables (/ga/ and /fa/) showed suppression. Results are discussed in terms of differences in saliency for acoustic versus visual features and which are more influenced by the social priming paradigm used.

5. **An instance-based account of mental arithmetic performance**

Evan Curtis, Randal Jamieson
University of Manitoba
evan.thomas.curtis@gmail.com, Randy.Jamieson@ad.umanitoba.ca

Simple arithmetic problems are often solved via direct fact retrieval. As such, an adequate model of mental arithmetic should be grounded in theories of human memory. However, current models of mental arithmetic are highly domain-specific and fail to draw upon general principles of memory. We present an instance-based account of mental arithmetic adapted from MINERVA2, a multitrace model of human memory. When presented with verification problems (e.g., $3 \times 7 = 21$, true or false?), the model responds most strongly to the correct solution, and also responds more strongly to table-related problems (e.g., 14, 28, 12, 18) than to unrelated problems. Humans demonstrate the same pattern of behaviour. The model does not possess a representation of the multiplication table; yet, it acts as if it does. We argue that aspects of mental arithmetic can be predicted by an interaction between the representation of arithmetic facts and basic processes of memory.

6. **A computer tool to simulate and teach associative learning**

Randall Jamieson¹, Matthew Crump², Samuel Hannah³
University of Manitoba¹, Brooklyn College, CUNY², University of Saskatchewan³
randy.jamieson@ad.umanitoba.ca, mcrump@brooklyn.cuny.edu, hannah.sam@gmail.com

We present a computer tool that simulates associative learning from three models of associative learning: the Rescorla and Wagner model (1972), the elemental model of Ghirlanda (2005), and the revised Rescorla and Wagner model of Van Hamme and Wasserman (1994). The tool is cross-platform and has an intuitive interface. Users specify a learning protocol in standard notation (e.g., blocking is encoded directly as “Training: A+, AB+; Test: A+, B+”) and specify model parameters within allowable ranges. Simulation results are reported in either a figure or table. We present the tool, demonstrate its use, and verify its accuracy against published simulations. Researchers can use the tool to model

learning data, compare the models, and predict results from novel protocols. Instructors can use the tool to teach computational analysis of associative learning in the classroom.

7. **Investigating the dynamics of search and decision making**

Don Zhang, Richard Anderson
Bowling Green State University
zdon@bgsu.edu, randers@bgsu.edu

Search is essential for human intelligence. Recent research on human search behavior has focused on the exploration exploitation trade-off and various heuristics used to model stopping behavior in search. For example, when search for mates, individuals have to decide whether to leave or stay in the relationship. In the current study, we create a two-stage search-decision task. Subjects first search in ambiguous search spaces to form choice sets; then, a decision is made from those choices. The environment is simulated in a task where subjects must assemble a team of cyclists for an upcoming race. The current study explores how people search in ambiguous environments and how their search process affects their decision-making. We explore various search heuristics and strategies people use. In doing so, we hope to better understand the underlying process and the dynamics of search and choice.

8. **Modeling the concept of “obscenity”**

Justin Miclat, Curt Burgess, Justin Estep
University of California, Riverside
jmich002@ucr.edu, curt@ucr.edu, jeste002@ucr.edu

The notion of obscenity is subjective as a result of the broad range of stimuli (whether films, periodicals and books, or photographs) that have been submitted for legal judgments. Former Supreme Court Justice Potter expressed this semantic quandary in 1973 when he deferred trying to define obscenity, but commented that “I know it when I see it.” *Miller v. California* (1973) determined that obscenity is comprised of three components: appeal to prurient interest, patent offensiveness, and literary value. Sets of words were generated using human judgments to represent the three components of obscenity and were then subjected to an analysis using the HAL model to test for conceptual independence. These sets of semantic concepts were then used to derive a neural network model of how the decision-making process could be modeled. This model offers a theoretical opportunity to begin investigating how decisions are made that involve obscenity.

9. **Formality prediction with data mining: Chinese political leaders as a case study**

Haiying Li, Arthur C. Graesser, Zhiqiang Cai
University of Memphis
haiyinglit@gmail.com, art.graesser@gmail.com, zhiqiang.cai@gmail.com

Data mining has the striking features of clustering and predicting. This study introduces a new method to predict and classify the formality of the political discourse with the data mining tool WEKA. First, we automatically analyzed the original Chinese political discourse excerpts produced by chairmen Mao Tse-tung, Deng Xiaoping and Jiang Zemin with Chinese Linguistic Inquiry Word Count, Chinese Latent Semantic Analysis, and Chinese Content Word Overlapping to obtain the value for each piece of discourse. Second, K-means clustering

and support vector machine (SVM) are used to develop unsupervised text mining. Thus, the excerpts are grouped into various clusters based on the frequencies of varied word categories. The results demonstrate that the predicting aspect estimates the likelihood for the formality of the political discourse. The best benefits of data mining are its ability to gain deeper understanding the previously unseen patterns with other empirical measures.

10. **Normative results for gender usage of emoticons: Reliability of meaning**

Jason Viray, Joanna Wong, Annie Chen, Curt Burgess, Sarah Maples

University of California, Riverside

jvira001@student.ucr.edu, jwong034@student.ucr.edu, achen039@student.ucr.edu,
curt@ucr.edu, smapl001@student.ucr.edu

Wolf (2000) completed a normative study of gender differences in the use of emoticons using text from several discussion groups. Although gender differences were found with some emoticons, the text sample (and, thus, the number of emoticons studied) was rather small (< 50,000 words). Furthermore, it was apparent that there was more variability that suspected in the meaning of even common emoticons. For example, the classic smiley face is understood to convey that one is happy. However, it was also used to convey joy. Yuki (2005) has shown that there are cultural differences in which emoticons tend to be used and in the meaning of emoticons. We report the results of a norming study where we used a large text corpus to determine relative usage and then asked male and female subjects to express their understanding of the meaning.

11. **Comparing fTMS and fTCD for studies of functional cerebral asymmetries**

David A. Washburn¹, Natasha Schultz¹, Johathan P. Gullledge², Holly A. Phillips¹

Georgia State University¹, Dalton State College²

dwashburn@gsu.edu, nbarrett78@yahoo.com, jgullledge@dsc.edu, hadams7@student.gsu.edu

Technological advances have brought a range of options to researchers who are interested in functional differences or specializations of the cerebral hemispheres. Behavioral paradigms (such as divided visual field tests) can be complemented with the information from fMRI, EEG, or other techniques. Two relatively recent innovations provide noninvasive and continuous assays of brain activity. Transcranial magnetic stimulation (TMS) uses rapid, and sometimes repeated, discharge of a magnetic coil to stimulate neural activity in an adjacent region. Researchers can use TMS to create temporary functional “lesions” by disrupting task-related neural firing in some area of the brain. In contrast, transcranial Doppler sonography (TCD) measures task-related changes in bloodflow velocity in the brain using ultrasound. We compared these two techniques in studies of functional cerebral asymmetries in the processing of numeric information. Relative strengths and weaknesses of each paradigm will be discussed.

12. **Garden path sentences between Han and Indigene: an fMRI study**

Yu-Chen Chan¹, Hsueh-Chih Chen¹, Yu-Lin Chang¹, Kuo-En Chang²

Department of Educational Psychology and Counseling, National Taiwan Normal University

¹, Graduate Institute of Information and Computer Education, National Taiwan Normal

University²-yccpsy@gmail.com, chcjyh@ntnu.edu.tw, gtyulin@gmail.com,
president@deps.ntnu.edu.tw

Previous imaging research has sought to identify the neural substrates of segregating the comprehension and elaboration processing by comparing garden path and general sentences (Chan et al., 2012). In this study, an attempt is made to compare the neural substrates of garden path sentences in Han and aboriginal groups via an event-related fMRI design. Our findings suggest that the IFG and SFG are responsible for incongruity detection, semantic decoding, semantic selection and semantic integration. Important between-ethnic differences also emerge, offering insight into disparate modes of language processing. One of these regions is the left inferior temporal gyrus (BA27), suggesting greater emphasis on language and executive processing in Han. Comparing responses to the Han and aboriginal group have been shown greater right IPL activity in Han during semantic processing of garden path sentences.

13. **CRIE: A tool for analyzing Chinese text characteristics**

Ju-Ling Chen, Jih-Ho Cha, Tao-Hsing Chang, Yao-Ting Sung, Kuan-Sheng Hsieh
Center of Learning Technology for Chinese Language, National Taiwan Normal University
Department of Computer Science and Information Engineering, National Kaohsiung
University of Applied Sciences
chenjuling@ntnu.edu.tw, toddcha@gmail.com, thchang@cs.nctu.edu.tw, sungtc@ntnu.edu.tw,
kuansheng@ntnu.edu.tw

Studies focusing on language analyses of alphabetic writing systems have been researched for more than 40 years. Abundant and advanced research outcomes and practical needs have made automated text analyzing tools possible; however, such tools designed for Chinese writing system are insufficient and scarce. Following the latest trend of multi-level analyses of text features (Graesser et al., 2004), we develop a tool called Chinese Readability Indices Explorer, CRIE, which can extract 90 features based on features of Chinese characters, words, syntax, and cohesion. The modules used in CRIE include lexicons, segmentation, syntactic parsers, corpora, latent semantic analysis, and other components that are widely used in computational linguistics. Not only does CRIE provide multi-level linguistic feature analyses, CRIE is also able to deal with literary Chinese and domain-specific texts. CRIE provides outputs on measures of individual linguistic features as well as providing formulas for different text domains , age groups.

14. **A Chinese word segmentation and POS tagging system for readability research**

Tao-Hsing Chang, Yao-Ting Sung, Yao-Tun Lee, Kuan-Sheng Hsieh
Department of Computer Science and Information Engineering, National Kaohsiung
University of Applied Sciences
Center of Learning Technology for Chinese Language, National Taiwan Normal University
thchang@cs.nctu.edu.tw, sungtc@ntnu.edu.tw, ddt0320@gmail.com,
kuansheng@ntnu.edu.tw

In recent years, readability research has relied on applications of natural language processing techniques to analyze documents. However, Chinese sentences consist of characters and with no blanks between words. Therefore, a mistake on word segmentation and/or part-of-speech tagging for Chinese sentences will result in many errors in the follow-up analysis. CRF model is recently the most popular and successful method for Chinese word segmentation. However, due to such problems as reiterative locution, unknown words and incomplete sentences, many readings for children cannot be processed accurately by CRF model. This

study aims to develop a Chinese word segmentation and POS tagging system called WeCan. This system is composed of bigram model, SPLR algorithm, unknown words extraction and rule bases. WeCan has been applied to the preprocessing procedure of CRIE. In preliminary experiments, it also worked well on the elementary school textbook in Taiwan.

15. **Constructing a Chinese text readability formula with multi-level linguistic features**

Jia-Fei Hong, Hou-Chiang Tseng, Yi-Shang Li, Yao-Ting Sung
Center of Learning Technology for Chinese Language, National Taiwan Normal University
jiafeihong@gmail.com, ouartz99@hotmail.com, bill.net.tw@yahoo.com.tw,
sungtc@ntnu.edu.tw

Previous readability research have adopted shallow linguistic features, which cannot fully reflect the complex process of reading comprehension. Given the differences between alphabetic and Chinese writing systems, the current study aims to select adequate features for Chinese texts, and construct a classifying model for Chinese texts. This study adopts 34 linguistic features under 10 dimensions selected from 4 levels, including word, semantics, syntax, and cohesion. Traditional linear and non-linear analyses did not address or compare the performances of uni-dimensional and multi-dimensional linguistic features. We therefore adopt discriminant analyses (DA) for linear analyses and the Support Vector Machine (SVM) for non-linear analyses to construct readability formulas, with both uni-dimensional and multi-dimensional features. By comparing the four approaches, this study shows that with multi-dimensional linguistic features, SVM can construct a relatively better mathematical readability model with an accuracy rate of texts classification reaching 69.95%.

16. **Integrating Cloze tests with Item Response Theory to validate a readability formulae**

Wei-Chun Lin, Ju-Ling Chen, Yao-Ting Sung
Center of Learning Technology for Chinese Language, National Taiwan Normal University
r93227107@ntu.edu.tw, chenjuling@ntnu.edu.tw, sungtc@ntnu.edu.tw

The present study validates a non-linear readability formula (Sung et al., 2012) with readers' performance in cloze tests, which integrates the item response theory to improve the deficits of traditional cloze tests. The first study adopts a two-factor mixed design to examine the performances of readers at different levels. The second study transforms the original cloze test scores through the Rasch Model with the purpose of evaluating text difficulty and the readers' reading ability, which improves not only the scoring of cloze tests but also the validity of cloze test measuring text difficulty. The results show that as compared with traditional cloze tests, the IRT-based cloze tests are superior in discriminant validity. Furthermore, the cloze tests scores show that the Sung et al. (2012) readability formulae can appropriately reflect the difficulty levels of texts.

17. **Evaluation of the feasibility of online readability application**

Hou-Chiang Tseng, Tao-Hsing Chang, Yao-Ting Sung
Center of Learning Technology for Chinese Language, National Taiwan Normal University
Department of Computer Science and Information Engineering, National Kaohsiung
University of Applied Sciences
ouartz99@hotmail.com, thchang@cs.nctu.edu.tw, sungtc@ntnu.edu.tw

Readability refers to the extent to which a text can be understood. Reading materials of high readability can facilitate reading comprehension, learning, and information retention. With the development of Cloud computing, it has become important to assess the readability of online texts. However, previous readability research focuses only on the accuracy of text classification but fails to comprehensively take into account the formulae accuracy, time complexity, and feature diversity. Currently readability formulae used in evaluating online documents are time-consuming and low in accuracy. Therefore, to enhance the online utilization of readability formulae, we propose a standardized framework for evaluating the feasibility of online readability formulae, based on the three major factors: (a) user waiting time, (b) text classification accuracy, and (c) feature diversity. We also verified the feasibility of this framework by comparing several readability models under such framework.

18. **Helping women decide about testing for genetic breast cancer risk with an Intelligent Tutoring System**

Christopher Wolfe¹, Valerie Reyna², Elizabeth Cedillos¹, Colin Widmer¹, Christopher Fisher¹, Priscila Brust-Renck²
Miami University¹, Cornell University²
wolfecr@muohio.edu, cedillem@muohio.edu, widmercl@muohio.edu,
fisherc2@muohio.edu, vr53@cornell.edu, pgb56@cornell.edu

We developed and tested the effectiveness of a web-based Intelligent Tutoring System (ITS) that engages women in a tutorial dialogue to help them understand and make decisions about genetic testing for breast cancer risk. This tutor, guided by fuzzy-trace theory, appears to be the first use of an ITS in medical decision-making. Using "expectations texts" and Latent Semantic Analysis, avatars engage women in dialogue about questions such as, "what should someone do if she finds out that she has inherited an altered BRCA gene?". The effectiveness of the ITS was tested in a randomized, controlled experiment. Participants were then given two tests of declarative knowledge about breast cancer and genetic risk, and twelve scenarios applying their knowledge. The ITS group scored significantly higher on two tests of declarative knowledge, with large effect sizes, and was also significantly better in applying their knowledge to scenarios to distinguish among risk levels.

19. **Exploring the effectiveness of a novel feedback mechanism within an intelligent tutoring system**

Jamal Williams, Xiangen Hu
Institute for Intelligent Systems, University of Memphis
jamalawilliams@gmail.com, xhu@memphis.edu

It has been shown that under the correct circumstances feedback can be an effective means of increasing student learning (Crooks, 1988; Black and Wiliam, 1998). According to some research, for feedback to be truly effective, a certain criteria must be met (Sadler, 2010). Unfortunately, in order to provide the previously mentioned criteria, teachers must devote a massive amount of time and resources to every individual student which is unrealistic. The Learner's Characteristic Curve (LCC) was implemented into an intelligent tutoring system known as AutoTutor Lite to simulate human tutoring feedback strategies. In a preliminary analysis of recent experiments, we found a significant difference on performance measures between conditions but not between groups. This might be due to the low sample size which

we hope to increase in future experiments. Furthermore, this finding might be a decent indicator of the LCC as an intelligent tutoring system feedback implementation.

20. **Training the cognitive benefits of video games: Can massively multiplayer online role playing games also improve performance?**

Emily Collins, Jonathan Freeman
Goldsmiths College, University of London
e.collins@gold.ac.uk, j.freeman@gold.ac.uk

Action video game players and trained non-players have demonstrated superior performance in a number of cognitive tasks (e.g. Green & Bavelier, 2006). However few have addressed the possibility of other games such as Massively Multiplayer Online Role Playing Games (MMORPGs) producing similar benefits. A total of 33 participants aged 18 to 29 were randomly allocated to play an action video game (Medal of Honor), MMORPG (Lord of the Rings Online) or a control game (Angry Birds) for a total of 10 hours over a three week period. Task switching, mental rotation, flanker compatibility, enumeration and visual short term memory performance were compared before and after the training period. An ANOVA revealed no significant effect of condition, suggesting that neither the action game nor the MMORPG improved performance more than the control game. Possible reasons for this lack of concordance with previous literature and the potential implications are discussed.

21. **Shyness, positive and negative affect, life satisfaction and video games**

Chris Koch¹, Mary Kate Koch²
George Fox University¹, Gonzaga University²
ckoch@georgefox.edu, mkoch@zagmail.gonzaga.edu

Although the relationship between shyness and social media use has been a target of research, the literature regarding shyness and video gaming is underdeveloped. In the present study, 127 (39 male and 88 female, Mage = 19.70) college students completed the Positive Affect and Negative Affect Scales (Watson, Clark, & Tellegen, 1988), the Shyness Scale (Cheek & Melichor, 1985), the Satisfaction with Life Scale (Diender, Emmons, Larsen, & Griffin, 1985), along with a demographic questionnaire. Respondents who missed a meal were more shy and expressed greater negative affect than those who had not. Respondents who missed a class or had been late for an appointment due to playing video games also had greater negative affect. These results suggest that negative affect, not shyness, is associated with disruptive gaming behaviors.

22. **End user development in psychology teaching**

Joerg Beringer
BeriSoft Cooperation
joerg.beringer@berisoft.com

Experiment control software vendors are in functional conflict between supporting the development of new lab experiments as well as enabling psychology students to design their own experiments despite a lack of computer skills. The latter design goal is called “end-user development” which refers to the concept of enabling domain experts that have no IT-skills to create and modify IT-artifacts (Henry Lieberman, Fabio Paternó, Markus Klann and Volker Wulf, End-User Development: an Emerging Paradigm, 2005). Depending on what target

audience the main focus is, commercial software packages take different routes to cater their end users: professional coding by technology savvy students vs. out-of-the box experiments downloaded from the web or prepared by power users for normal students. This presentation brings together various established vendors to discuss the trade-offs between standard programming languages, domain-specific languages, and re-usable content to discuss the different strategies and design rationales of empowering students to become active researchers.

23. **Examining the modality effect in multimedia self-paced instruction**

Dmitrii Paniukov, Ismahan Arslan-Ari, Steven Crooks

Texas Tech University

dmitry.panyukov@ttu.edu, ismahan.arslan@ttu.edu, steven.crooks@ttu.edu

The purpose of this study was to investigate the effects of modality (written text vs. narrated text) on the learning and mental effort of participants studying a computer-based static diagram at their own pace. In Experiment 1, participants were randomly assigned to four versions of the computer-based materials, forming four counterbalanced conditions in a within subject design. The results of a transfer test approached significance evidencing a reverse modality effect ($F(3, 43) = 3.953, p = .053$, wherein participants studying written text outperformed those studying narrated text. However, a matching test revealed a modality effect ($F(3, 43) = 4.686, p = .036$), wherein participants studying narrated text outperformed those studying written text. During September-October 2012 another experiment will collect data with an Eye tracking device to investigate why both modality and reverse modality effects exist in the same instruction. The results will be presented at the conference.

24. **A redundancy effect in self-paced multimedia instruction**

Ismahan Arslan-Ari, Dmitrii Paniukov, Steven Crooks

Texas Tech University

ismahan.arslan@ttu.edu, dmitry.panyukov@ttu.edu, steven.crooks@ttu.edu

The purpose of this study is to investigate the effects of redundancy (written text vs. written and narrated text) on the learning and mental effort of participants studying a computer-based static diagram at their own pace. In Experiment 1 participants were randomly assigned to four versions of the computer-based materials, forming four counterbalanced conditions in a within subject design. The results of a matching test revealed a reverse-redundancy effect, wherein participants studying written and narrated text outperformed those studying written text only ($F(3, 16) = 4.284, p = .056$). During September-October 2012 an additional experiment will be conducted with an Eye tracking device to investigate the cause of the reverse redundancy effect in self-paced multimedia instruction. The results will be presented at the conference.

25. **Unified data repository for psychology experiments**

Yuchen Liao, Xudong Huang, Jun Xie, Xiangen Hu

The University of Memphis

yliao2@memphis.edu, xhuang3@memphis.edu, jxie2@memphis.edu, xhu@memphis.edu

This project is an online database system which enables the users to efficiently organize all the data from the psychology experiments and to retrieve and exchange information from the data

collected from experiments. This system would enhance a researcher's ability to keep track of discoveries and share information with others. Primarily any researcher who conducts experiments can store the project information in this database. The information retrieved from the projects when kept together in the same database may produce new results which can be used further. The researchers could also retrieve data which is collected long ago by this system. For the long term benefit, there has to be a way to combine and analyze data from different experiments. So, there is need for a database which can store data related to different experiments with different experimental designs, a database from which data can be retrieved after many years.

26. **EHME: A new word database for research in Basque language**

Joana Acha, Pello Salaburu, Itziar Laka

University of the Basque Country

joana.acha@ehu.es, pello.salaburu@ehu.es, itziar.laka@ehu.es

We will describe an online program that enables students and researchers on psycholinguistics to extract word stimuli, based on a broad range of statistics concerning the properties of words and nonwords in (Basque). This program has been recently developed to overcome some limitations of the one that was available up to this moment, due to the increasing demand of research on Basque language. This new programs includes a greater number of words and sources, and includes measures of neighborhood frequency (taking into account recent evidence about transposed, addition and deletion neighbors; Davis et al., 2009) and morphological structure frequency apart from the classical word frequency (at the whole-word and lemma levels), bigram and biphone frequency, orthographic similarity, orthographic and phonological structure, and syllable-based measures. It is designed for use by researchers in psycholinguistics, particularly those concerned with recognition of isolated words and morphology. In addition the program can be used to extract words from statistical criteria (<http://www.ehu.es/ehg/ehme/datu2hitz.htm>), as well as to obtain statistical characteristics form a list of words (<http://www.ehu.es/ehg/ehme/hitz2datu.htm>).

27. **PsyToolkit: Using the web to program and share experiments**

Gijsbert Stoet

University of Leeds

g.stoet@leeds.ac.uk

There are many open-source software packages for programming and running experiments, but few run on all computer platforms. There are many advantages of experiments that can run on all platforms. It eases sharing of experiments, and it is eases teaching. To provide this functionality, there is now a web-based PsyToolkit version. Free of cost, users can create a PsyToolkit account via psytoolkit.leeds.ac.uk. Once logged in, there are tutorials and examples. Further, experiments can be edited and run (as Java applications) from within a web browser. There are many advantages in terms of sharing and demonstrating cognitive experiments. The web based version has special options for psychology teachers. Note 1: The measurement of response time is not as precise as when run as a Linux application, even though the end-user does not notice any difference. Note 2: The current version works, but is new and still under development."

28. **Libet's clock goes online: Using Labclock and Labclock Web to conduct offline and online studies on conscious will**

Pablo Garaizar, Carmelo Perez-Cubillas, Helena Matute
Universidad de Deusto, Bilbao, Spain
garaizar@deusto.es, carmelo.perez@deusto.es, matute@deusto.es

The role of consciousness in decision making is an important research field in psychology. Most of the studies about it are conducted using Libet's clock paradigm. However, the technical details of the apparatus used to conduct this kind of experiments are rarely explained. For this reason, we developed Labclock, a Microsoft Windows program - provided with multimedia timers and high resolution timing functions - that implements Libet's clock paradigm. Similarly, Labclock Web is an online version of Labclock that uses HTML5-related APIs to accomplish required timing accuracy. Both versions have been tested with very promising results. Our main goal is to provide a free implementation of Libet's clock paradigm that will foster new experiments on conscious will.

29. **An Android app for the Nonverbal Stroop Card Sorting Task**

Chris Koch, Josiah Hotovec
George Fox University
ckoch@georgefox.edu, jhotovec09@georgefox.edu

The Nonverbal Stroop Card Sorting Task (NSCST, Koch & Roid, 2012) is a nonverbal Stroop assessment that requires examinees to sort cards according to target colors. Two color blocks appear on each card with a cross being centered in block to indicate that it is the target. In the congruent condition, both color blocks are the same color. In the incongruent condition, the two blocks are different colors. Sorting times are longer for the incongruent condition than for the congruent condition. Although the task is relatively straightforward and easy to administer, the use cards does place some constraints on where and how the test can be administered. Therefore, an Android app was developed to administer the NSCST using a tablet or other Android device. Data comparing the physical and app versions of the task was collected. Implications for testing are discussed.

30. **S-Delta plots; a method to display RT distributions across multiple subjects**

Pablo Gomez, Robert Zimmerman
DePaul University
pgomez1@depaul.edu, rzimmer1@depaul.edu

Researchers explore if variables facilitate or inhibit processing. Often, these issues are examined using averaged RTs, which do not allow us to explore the distributional properties of the data. We propose a method (related to delta-plots) to visualize the across-subject variability and the RT distributions simultaneously. For each participant we calculate the RT at the .1, .5 and the .75 quantiles. Once these three RTs are calculated for each subject for each condition, residuals can be obtained. Subjects are sorted based on the residual .1q and these numbers are plotted based on the sorting creating a monotonically growing function. Using the order obtained with the .1q, we then plot smoothed lines based on the residual .5q and .75q. We termed this graph a Delta-S plot: an EDA method that we propose researchers

should use when analyzing RT data from psychology experiments because different mechanisms have different Delta-S signatures.

31. **Preliminary results of new reaction time tasks of detecting velocity change of a moving stimulus to assess visual attention skills in Occupational Therapy**

Toshio Ohyanagi¹, Kunihiro Kanaya¹, Yasuhito Sengoku¹, Masako Miyazaki²
Sapporo Medical University¹, University of Alberta²
ohyanagi@sapmed.ac.jp, kanayak@sapmed.ac.jp, sengoku@sapmed.ac.jp,
mmiyazak@ualberta.ca

Visual attention is essential for safely executing any activities that include activities of daily livings. The assessment of visual attention skills is one of the important roles in Occupational Therapy. Occupational Therapists have been employing neuropsychological tests and/or reaction time (RT) tasks to conduct the assessment. However it was often the case that the results did not reflect patient's behavior identified by an experienced Occupational Therapist. We developed new RT tasks that are to detect a stimulus of changing its velocity among thirty-six stimuli that move to the same direction at the same velocity. The stimuli are circles of one degree of visual angle in diameter and displayed on a screen subtended 46 degrees of visual angle horizontally and 26 degrees vertically. We will present our preliminary results of the assessment of visual attention skills using the tasks. We would like to share how well the results reflect subjects' behavior.

32. **Could your equipment account for your experimental effect?**

Richard Plant¹, Garry Turner²
University of Sheffield, UK¹, University of York, UK²
r.plant@blackboxtoolkit.com, garry.turner@york.ac.uk

Widespread use of modern technologies has left many assuming they no longer need be concerned with the intricacies of millisecond presentation, synchronisation and response timing. However, through empirical investigation, we have discovered numerous sources of timing error within live studies (e.g. Plant et al., 2009). Timing errors can have many causes and are not wholly predictable as to what may, or may not, affect timing within certain types of paradigm. To more practically aid the researcher, we have developed a method for benchmarking timing in the majority of paradigms whilst running in-situ and without modification on the researchers own hardware. In this paper we introduce our new ARM Cortex based benchmarking hardware, stress the importance of such independent validation for Psychology as a whole, and highlight typical areas that can be subject to error. Case studies from the fields of vision, judgment and decision making and space research are presented.

.....**Notes**.....

.....Notes.....